

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Hirofumi MURATANI

GAU:

SERIAL NO: New Application

EXAMINER:

FILED: Herewith

FOR: DIGITAL WATERMARK EMBEDDING APPARATUS AND METHOD, AND DIGITAL
WATERMARK ANALYSIS APPARATUS, METHOD AND PROGRAM

REQUEST FOR PRIORITY

COMMISSIONER FOR PATENTS
ALEXANDRIA, VIRGINIA 22313

SIR:

- Full benefit of the filing date of U.S. Application Serial Number , filed , is claimed pursuant to the provisions of 35 U.S.C. §120.
- Full benefit of the filing date(s) of U.S. Provisional Application(s) is claimed pursuant to the provisions of 35 U.S.C. §119(e): Application No. Date Filed
- Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

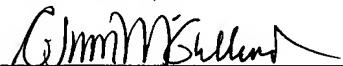
<u>COUNTRY</u>	<u>APPLICATION NUMBER</u>	<u>MONTH/DAY/YEAR</u>
Japan	2002-248941	August 28, 2002

Certified copies of the corresponding Convention Application(s)

- are submitted herewith
- will be submitted prior to payment of the Final Fee
- were filed in prior application Serial No. filed
- were submitted to the International Bureau in PCT Application Number
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and
- (B) Application Serial No.(s)
 are submitted herewith
 will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.



Marvin J. Spivak

Registration No. 24,913

C. Irvin McClelland
Registration Number 21,124



22850

Tel. (703) 413-3000
Fax. (703) 413-2220
(OSMMN 05/03)

0350098

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office

出願年月日

Date of Application: 2002年 8月28日

出願番号

Application Number: 特願2002-248941

[ST.10/C]:

[JP2002-248941]

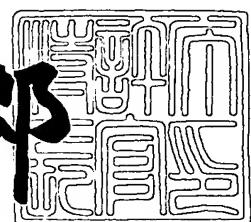
出願人

Applicant(s): 株式会社東芝

2003年 1月31日

特許庁長官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3002794

【書類名】 特許願
【整理番号】 A000202036
【提出日】 平成14年 8月28日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 15/00
【発明の名称】 電子透かし埋め込み装置、電子透かし解析装置、電子透かし埋め込み方法、電子透かし解析方法及びプログラム
【請求項の数】 26
【発明者】
【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内
【氏名】 村谷 博文
【特許出願人】
【識別番号】 000003078
【氏名又は名称】 株式会社 東芝
【代理人】
【識別番号】 100058479
【弁理士】
【氏名又は名称】 鈴江 武彦
【電話番号】 03-3502-3181
【選任した代理人】
【識別番号】 100084618
【弁理士】
【氏名又は名称】 村松 貞男
【選任した代理人】
【識別番号】 100068814
【弁理士】
【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 電子透かし埋め込み装置、電子透かし解析装置、電子透かし埋め込み方法、電子透かし解析方法及びプログラム

【特許請求の範囲】

【請求項1】

デジタルコンテンツの複製物に電子透かしとして埋め込むべき識別情報に一意に対応する、序列を持った複数のシンボルの並びからなるシンボル列を生成する第1の生成手段と、

生成された前記シンボル列を構成する個々の序列のシンボルごとに、対応する埋め込み符号をそれぞれ生成する第2の生成手段と、

生成された前記個々の序列のシンボルにそれぞれ対応する埋め込み符号を、前記デジタルコンテンツの複製物に電子透かしとして埋め込む符号埋め込み手段とを備えたことを特徴とする電子透かし埋め込み装置。

【請求項2】

前記第2の生成手段は、前記シンボル列の個々の序列において、同一の序列に属する互いに異なるシンボルに対応する埋め込み符号同士が、相互相關の無い又は小さい乱数列になるように、前記埋め込み符号を生成することを特徴とする請求項1に記載の電子透かし埋め込み装置。

【請求項3】

前記第2の生成手段は、前記シンボル列の異なる序列に属するシンボルに対応する埋め込み符号同士についても、相互相關の無い又は小さい乱数列になるよう、前記埋め込み符号を生成することを特徴とする請求項2に記載の電子透かし埋め込み装置。

【請求項4】

前記識別情報は、整数値であり、

前記シンボル列の各序列ごとに対応して特定の整数値が予め定められており、

前記第1の生成手段は、前記識別情報を前記シンボル列の個々の序列に対応する前記整数値でそれぞれ除し、これによって得られた個々の序列に対応する剰余値を当該序列に従って並べたものを、生成すべきシンボル列とすることを特徴と

する請求項1に記載の電子透かし埋め込み装置。

【請求項5】

前記識別情報は、整数値であり、

前記シンボル列の各序列ごとに対応して特定の整数値が予め定められており、

前記第1の生成手段は、前記識別情報を前記シンボル列の個々の序列に対応する前記整数値でそれぞれ除し、これによって得られた個々の序列に対応する剰余値をそれぞれ一意に対応する情報に変換し、これによって得られた個々の序列に対応する情報を当該序列に従って並べたものを、生成すべきシンボル列とすることを特徴とする請求項1に記載の電子透かし埋め込み装置。

【請求項6】

前記シンボル列の各序列ごとに対応して予め定められた特定の整数値は、互いに素の関係にある整数値であることを特徴とする請求項4または5に記載の電子透かし埋め込み装置。

【請求項7】

前記デジタルコンテンツの複製物に電子透かしとして埋め込むために使用可能とする前記識別情報の範囲を、前記シンボル列を割り当て可能な識別情報の範囲よりも狭い範囲に制限したことを特徴とする請求項1ないし6のいずれか1項に記載の識別情報復号装置。

【請求項8】

前記識別情報は、前記デジタルコンテンツの複製物を提供するユーザに固有の情報を有することを特徴とする請求項1ないし7のいずれか1項に記載の電子透かし埋め込み装置。

【請求項9】

互いに異なる識別情報を電子透かしとして埋め込まれた複数のデジタルコンテンツの複製物をもとにして結託攻撃によって作出された不正のデジタルコンテンツの複製物を対象として、結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定するための電子透かし解析装置であって、

前記デジタルコンテンツの複製物から、序列を持った複数の埋め込み符号をそれぞれ抽出する第1の抽出手段と、

抽出された前記複数の埋め込み符号にそれぞれ対応するシンボルを求め、求められた複数のシンボル列を、対応する埋め込み符号の序列に従って並べてシンボル列を求める第2の抽出手段と、

前記結託攻撃に用いられたデジタルコンテンツの複製物から抽出された前記シンボル列及び各々の正当な識別情報に対してそれぞれ一意に割り当てられるシンボル列に基づいて、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定する追跡手段とを備えたことを特徴とする電子透かし解析装置。

【請求項10】

結託攻撃を受ける前の前記デジタルコンテンツの複製物に電子透かしとして埋め込む前記識別情報には、当該識別情報に一意に対応する、序列を持った複数のシンボルの並びからなるシンボル列が割り当てられており、

前記デジタルコンテンツの複製物には、前記シンボル列を構成する個々の序列のシンボルごとにそれぞれ対応する埋め込み符号が電子透かしとして埋め込まれていることを特徴とする請求項9に記載の電子透かし解析装置。

【請求項11】

前記第1の抽出手段は、個々の前記序列ごとに、前記結託攻撃において最も多く使用されたと評価される埋め込み符号を抽出することを特徴とする特徴とする請求項9または10に記載の電子透かし解析装置。

【請求項12】

前記埋め込み符号は、前記シンボル列の個々の序列において、同一の序列に属する互いに異なるシンボルに対応する埋め込み符号同士が、相互相関の無い又は小さい乱数列になるように、生成されたものであり、

前記第1の抽出手段は、個々の前記序列ごとに、当該序列で使用される複数の埋め込み符号の各々について、前記デジタルコンテンツの複製物との相関を取り、相関値が最大であった埋め込み符号を抽出することを特徴とする請求項9ないし11のいずれか1項に記載の電子透かし解析装置。

【請求項13】

前記埋め込み符号は、前記シンボル列の異なる序列に属するシンボルに対応す

る埋め込み符号同士についても、相互相關の無い又は小さい乱数列になるように、生成されたものであることを特徴とする請求項12に記載の電子透かし解析装置。

【請求項14】

前記第2の抽出手段は、全ての識別情報について、当該識別情報に一意に対応するシンボル列と、前記対象となったデジタルコンテンツの複製物から求められた抽出シンボル列とを比較した結果に基づいて、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定することを特徴とする請求項9ないし13のいずれか1項に記載の電子透かし解析装置。

【請求項15】

前記第2の抽出手段は、全ての識別情報について、当該識別情報に一意に対応するシンボル列と、前記対象となったデジタルコンテンツの複製物から求められた抽出シンボル列とを、各序列ごとにそれぞれ比較し、予め定められた個数以上の序列におけるシンボルの一致が得られた場合に、当該識別情報を、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報として特定することを特徴とする請求項14に記載の電子透かし解析装置。

【請求項16】

前記識別情報は、整数値であり、且つ、前記シンボル列の各序列 i ごとに對応して特定の整数値 $N(i)$ が予め定められており（ここで、 $N(0) \leq N(1) \leq \dots \leq N(M)$ とする）、且つ、前記第1の生成手段は、前記識別情報の値に基づき、前記シンボル列の個々の序列 i に対応するシンボル $S(i)$ を、0から $N(i) - 1$ の範囲で選択し、これによって得られた個々の序列 i に対応するシンボル $S(i)$ を当該序列に従って並べたものを、生成すべきシンボル列とするものである場合に、

前記予め定められた個数を、 $k + 1$ とする（ただし、 k は、 $N(0), \dots, N(k)$ の積が前記識別情報の総数以上となるような値であり、1は $[1 - \prod_{i=0}^{k-1} N(i)]^S \geq 1 - \varepsilon_2$ （ただし、 \prod をとる i の範囲は、 $i = 0 \sim (k - 1)$ 又は $i = k \sim (k + 1 - 1)$ であり、 S は、 $M C_{k+1}$ であり、 $0 < \varepsilon_2 < 1$ ）は、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込

まれていた識別情報として誤った識別情報を特定してしまう追跡誤り率である)により与えられる値である)ことを特徴とする請求項15に記載の電子透かし解析装置。

【請求項17】

前記識別情報は、整数値であり、且つ、前記シンボル列の各序列 i ごとに対応して特定の整数値 $N(i)$ が予め定められており(ここで、 $q = N(0) = N(1) = \dots = N(M)$ とする)、且つ、前記第1の生成手段は、前記識別情報の値に基づき、前記シンボル列の個々の序列 i に対応するシンボル $S(i)$ を、0から $N(i)-1$ の範囲で選択し、これによって得られた個々の序列 i に対応するシンボル $S(i)$ を当該序列に従って並べたものを、生成すべきシンボル列とするものである場合に、

前記予め定められた個数を、 $k+1$ とする(ただし、 k は、 $N(0), \dots, N(k)$ の積が前記識別情報の総数以上となるような値であり、1は $[1 - 1/q]^k \leq 1 - \varepsilon$ (ただし、 S は、 $\binom{M}{k+1}$ であり、 ε (ただし、 $0 < \varepsilon < 1$)は、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報として誤った識別情報を特定してしまう追跡誤り率である)により与えられる値である)ことを特徴とする請求項15に記載の電子透かし解析装置。

【請求項18】

前記第2の抽出手段は、予め定められた最大結託数以下の結託数による任意の識別情報の組合せからなる結託グループのうち、前記対象となったデジタルコンテンツの複製物から抽出された前記シンボル列を作出可能な1又は複数の結託グループを特定し、該特定した1又は複数の結託グループをそれぞれ構成する識別情報の組合せに基づいて、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報として特定することを特徴とする請求項9ないし13のいずれか1項に記載の電子透かし解析装置。

【請求項19】

前記第2の抽出手段は、前記対象となったデジタルコンテンツの複製物から抽出された前記シンボル列を作出可能な結託グループが1つのみ特定された場合には、該特定された結託グループを構成する識別情報を全て、前記結託攻撃に用い

られたデジタルコンテンツの複製物に埋め込まれていた識別情報として特定することを特徴とする請求項18に記載の電子透かし解析装置。

【請求項20】

前記第2の抽出手段は、前記対象となったデジタルコンテンツの複製物から抽出された前記シンボル列を作出可能な結託グループが複数特定された場合には、該特定された結託グループの全てに共通に含まれる識別情報のみを、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報として特定することを特徴とする請求項18または19に記載の電子透かし解析装置。

【請求項21】

前記デジタルコンテンツの複製物に電子透かしとして埋め込むために使用可能とする前記識別情報の範囲を、前記シンボル列を割り当て可能な識別情報の範囲よりも狭い範囲に制限したことを特徴とする請求項9ないし20のいずれか1項に記載の電子透かし解析装置。

【請求項22】

前記識別情報は、前記デジタルコンテンツの複製物を提供するユーザに固有の情報をあることを特徴とする請求項9ないし21のいずれか1項に記載の電子透かし解析装置。

【請求項23】

デジタルコンテンツの複製物に電子透かしとして埋め込むべき識別情報に一意に対応する、序列を持った複数のシンボルの並びからなるシンボル列を生成し、

生成された前記シンボル列を構成する個々の序列のシンボルごとに、対応する埋め込み符号をそれぞれ生成し、

生成された前記個々の序列のシンボルにそれぞれ対応する埋め込み符号を、前記デジタルコンテンツの複製物に電子透かしとして埋め込むことを特徴とする電子透かし埋め込み方法。

【請求項24】

互いに異なる識別情報を電子透かしとして埋め込まれた複数のデジタルコンテンツの複製物をもとにして結託攻撃によって作出された不正のデジタルコンテンツの複製物を対象として、結託攻撃に用いられたデジタルコンテンツの複製物に

埋め込まれていた識別情報を特定するための電子透かし解析方法あって、

前記デジタルコンテンツの複製物から、序列を持った複数の埋め込み符号をそれぞれ抽出し、

抽出された前記複数の埋め込み符号にそれぞれ対応するシンボルを求め、求められた複数のシンボル列を、対応する埋め込み符号の序列に従って並べてシンボル列を求め、

前記結託攻撃に用いられたデジタルコンテンツの複製物から抽出された前記シンボル列及び各々の正当な識別情報に対してそれぞれ一意に割り当てられるシンボル列に基づいて、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定することを特徴とする電子透かし解析方法。

【請求項25】

コンピュータを電子透かし埋め込み装置として機能させるためのプログラムであって、

デジタルコンテンツの複製物に電子透かしとして埋め込むべき識別情報に一意に対応する、序列を持った複数のシンボルの並びからなるシンボル列を生成する機能と、

生成された前記シンボル列を構成する個々の序列のシンボルごとに、対応する埋め込み符号をそれぞれ生成する機能と、

生成された前記個々の序列のシンボルにそれぞれ対応する埋め込み符号を、前記デジタルコンテンツの複製物に電子透かしとして埋め込む機能とをコンピュータに実現させるためのプログラム。

【請求項26】

互いに異なる識別情報を電子透かしとして埋め込まれた複数のデジタルコンテンツの複製物をもとにして結託攻撃によって作出された不正のデジタルコンテンツの複製物を対象として、結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定するための電子透かし解析装置として、コンピュータを機能させるためのプログラムであって、

前記デジタルコンテンツの複製物から、序列を持った複数の埋め込み符号をそれぞれ抽出する機能と、

抽出された前記複数の埋め込み符号にそれぞれ対応するシンボルを求め、求められた複数のシンボル列を、対応する埋め込み符号の序列に従って並べてシンボル列を求める機能と、

前記結託攻撃に用いられたデジタルコンテンツの複製物から抽出された前記シンボル列及び各々の正当な識別情報に対してそれぞれ一意に割り当てられるシンボル列に基づいて、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定する機能機能とをコンピュータに実現させるためのプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、識別情報に対応する埋め込み符号を生成しコンテンツの複製物に埋め込む電子透かし埋め込み装置及び電子透かし埋め込み方法、結託攻撃に用いられたコンテンツの複製物に埋め込まれていた埋め込み符号に対応する識別情報を求める電子透かし解析装置及び方法並びにプログラムに関する。

【0002】

【従来の技術】

デジタルコンテンツ（例えば、静止画、動画、音声、音楽等）は、多数のデジタルデータで構成された構造を持つ。そして、その構造の中には、データを変更しても、当該デジタルコンテンツの作品の同一性あるいは経済的価値を保持できる部分がある。そのような許容された範囲内のデータを適宜変更するによって、デジタルコンテンツに、種々の情報を埋め込むことができる。このような技術は、電子透かしとして、良く知られている。

【0003】

電子透かし技術によって、デジタルコンテンツに、様々な透かし情報（例えば、コンテンツの著作権者やユーザを識別する情報、著作権者の権利情報、コンテンツの利用条件、その利用時に必要な秘密情報、コピー制御情報等、あるいはそれらを組み合わせたものなど）を、様々な目的（例えば、利用制御、コピー制御を含む著作権保護、二次利用の促進等）で埋め込み、検出・利用することができ

る。

【0004】

ここでは、例えば同一のデジタルコンテンツを多数のユーザを対象として配給するときに適用される技術として、デジタルコンテンツの複製物に、当該複製物を個々に識別するための情報（例えば、ユーザIDに一意に対応する透かし情報）を埋め込む場合を考える。

【0005】

デジタルコンテンツの複製物に固有の識別情報（に対応する符号）を埋め込む手法は、そのデジタルコンテンツの複製物が更に複製されて海賊版として出回ったときに、該海賊版から識別情報を検出することによって流出元ユーザを特定することができることから、デジタルコンテンツの違法コピーに対する事前の抑制として機能するとともに、著作権侵害が発生したときの事後の救済にも役立つことになる。

【0006】

また、あるユーザがデジタルコンテンツの複製物に埋め込まれた識別情報を無効にするためには、ユーザにはどの部分が識別情報（に対応する符号）を構成するビットであるか分からないので、当該デジタルコンテンツの複製物に相当の改変を加える必要があり、そうすると、当該デジタルコンテンツの経済的価値を損なってしまうので、違法コピーの動機付けを奪うことができる。

【0007】

このような状況において違法コピーを可能ならしめる方法として出現したのが、「結託攻撃（collusion attack）」である。

【0008】

結託攻撃は、異なる複製物には異なる識別情報が埋めこまれていることを利用するものである。例えば、複数人で複製物を持ち寄って、それらをビット単位で比較することによって、デジタルデータの値が異なる部分を見つけ出し、その部分を改ざん（例えば、多数決、少数決、ランダマイズ等）することによって、識別情報を改ざん、消失させる。あるいは、具体的な比較操作は行わず、コンテンツ間で画素値を平均化するなどの操作を行って、識別情報を改ざん、消失させる

【0009】

例えば簡単な例で示すと、A氏、B氏、C氏の複製物にそれぞれ、

A : 1 0 … 0 0 …

B : 0 0 … 1 1 …

C : 1 1 … 0 1 …

という識別情報（実際には、これに対応する符号）が埋め込まれていた場合に、例えば、多数決あるいは平均化等によって、

1 0 … 0 1 …

というように、A氏、B氏、C氏のいずれとも異なる識別情報が検出されるようなコンテンツを出現させることができてしまう。

【0010】

そこで、結託攻撃に対する耐性、すなわち結託攻撃を受けても結託者の全部または一部を特定できるような性質を持つ符号（結託耐性符号と呼ばれる）を電子透かしとして埋め込む方法および該結託耐性符号に基づく追跡アルゴリズム（tracing algorithm；結託攻撃に用いられたコンテンツに埋め込まれた識別番号を特定し、結託者のユーザIDを特定するためのアルゴリズム）が種々提案されている。例えば、c-secure 符号(D. Boneh and J. Shaw, “Collusion-Secure Fingerprinting for Digital Data,” CRYPTO’ 95, 180-189, 1995.)やc-secure CRT 符号(Hirofumi Muratani, “A Collusion-Secure Fingerprinting Code Reduced by Chinese Remaindering and its Random-Error Resilience”, Information Hiding Workshop 2001, 301-315, 2001.や、特開2001-285623号公報)がその例である。

【0011】

ここで、c-secure CRT符号について簡単に説明する。

【0012】

c-secure CRT符号は、M個の成分符号を連接してなる連接符号の構造を持つ。例えば、第1番目の成分符号W(i)は、0のみからなるか、1のみからなるか、または0と1が混在する（ただし、0と1の境界はただ1つであ

る)、所定長のビット列であって、識別情報 u を法 p (i) で除したときの剰余によって、成分符号 W (i) に含まれる 0 のビット数及び 1 のビット数、すなわち、成分符号 W (i) を構成するビット列における 0 の系列と 1 の系列との境界位置が決定されるような符号である。

【0013】

例えば、 $M = 3$ とし、 $p(1) = 3$, $p(2) = 5$, $p(3) = 7$ とし、成分符号 W (i) として $\Gamma_0(n, d)$ 符号 (1 または 0 のみからなる連続したビットを一つの単位 B (j) とし、 $B(0) \sim B(n-2)$ を連結したもの; ただし、 $B(0) \sim B(n-2)$ は、すべてが 0 のみからなるか、すべてが 1 のみからなるか、または $B(0) \sim B(m-1)$ までは 0 のみからなり且つ $B(m) \sim B(n-2)$ までは 1 のみからなるもの) を用いるものとし ($d = 3$ とする)、識別情報 = ユーザ ID とすると、

例えば、ユーザ ID = 2 に対応する符号は、

000000 000000111111 000000111111111111

となり、ユーザ ID = 3 に対応する符号は、

111111 000000000111 000000000111111111

となる。

【0014】

この場合に、ユーザ ID = 2 のユーザとユーザ ID = 3 の 2 人のユーザが持ちよったコンテンツを比較すると、上記 36 ビットのうち、左から 1~6 番目、13~15 番目、25~27 番目が相違していることがわかる。そこで、それらが識別情報に対応する符号の一部と分かるため、1~6 番目、13~15 番目、25~27 番目のうちの一部に改ざんが施され、例えば、

010101 000000010111 000000101111111111

というように、ユーザ ID = 2 とユーザ ID = 3 のいずれとも異なる識別情報が検出されるようなコンテンツが作出される。

【0015】

ところが、正当なユーザ ID に対応する符号は、予め定められたブロックのサイズ d (上記の例の場合、3) ビット未満の数の 1 や 0 が孤立して存在する部分

は存在しないが、これに対して、上記のように結託攻撃を受けたコンテンツから検出される符号は、持ち寄ったコンテンツ同士で相違する符号部分については、0と1が混在しており、すなわち、 d （上記の例の場合、3）ビット未満の数の1や0が孤立して存在する部分が存在することが分かる。

【0016】

そこで、追跡アルゴリズムでは、検出された符号の各々の成分符号を調べ、予め定められた d （上記の例の場合、3）ビット未満の数の1や0が孤立して存在する成分符号が検出された場合に、結託攻撃がなされたものと判断することができる。

【0017】

ところで、c-secure CRT符号では、追跡対象コンテンツから検出された符号の第 i 番目の成分符号 $W(i)$ を、左端のビットからみていったときにはじめて出現する、0のみからなる要素と1を含む要素との境界の位置（を表す整数値）を最小剰余とし、右端のビットからみていったときにはじめて出現する、1のみからなる要素と0を含む要素との境界の位置（を表す整数値）を最大剰余とすると、最大剰余及び最小剰余（=剰余対）は、いずれかの結託者の識別情報 u に対する剰余すなわち $u \bmod p(i)$ と等しくなるという性質がある。なお、第 i 番目の成分符号 $W(i)$ が0のみからなる場合には、最小剰余=最大剰余= $p(i) - 1$ 、第 i 番目の成分符号 $W(i)$ が1のみからなる場合には、最小剰余=最大剰余=0、0のみからなる要素と1を含む要素との境界がない場合には、最小剰余=0、1のみからなる要素と0を含む要素との境界がない場合には、最大剰余= $p(i) - 1$ とする。

【0018】

したがって、追跡対象コンテンツから検出されたc-secure CRT符号の各成分部号から求めた剰余対を解析することによって、結託攻撃に使用された複製物に埋め込まれていたであろう識別情報の一部又は全部を求めることができるようになる。

【0019】

上記の例の場合、

010101 000000010111 000000101111111111

という符号から、

結託者の識別情報として、ユーザID=2とユーザID=3を特定することができる。

【0020】

このc-secure CRT符号は、次のようなマーキング仮定に基づいています。すなわち、結託攻撃に参加している結託者の中に、ある成分符号中のあるビットに対して異なる値を持つ者がいる場合に、結託攻撃の結果、生成される成分符号におけるそのビットがいずれの値をとるかは、確率的に決まる。よって、ブロックのサイズdがある程度大きくなると、正しく最小剩余及び最大剩余を検出することができる。

【0021】

【発明が解決しようとする課題】

c-secure CRT符号において、具体的の場合においては、追跡対象コンテンツから上記の最小剩余及び最大剩余を正しく検出できないケースが存在し得る。

【0022】

例えば、100人での結託攻撃を考えた場合、成分符号中あるブロックにおいて、99人にはビット列0が残りの一人にはビット列1が割り当てられていたとする。マーキング仮定では、0でも1でもないビット列がある確率で検出されることを期待している。ところが、電子透かしでは、ある測定量が所定のしきい値を越えるか否かでビット値の判定を行う場合が多く、結託攻撃がコンテンツの平均操作の場合、結託攻撃後の測定量は攻撃前の測定量の平均となると期待され、99対1の平均では、99人側の測定量の影響が大きく、ビット列は、99人側の値0として検出される確率が大きい。このブロックが最小剩余及び最大剩余を検出する境界のブロックならば、正しく最小剩余及び最大剩余を検出できないことになる。なお、これを符号感度の問題と呼ぶ。

【0023】

本発明は、上記事情を考慮してなされたもので、結託攻撃に用いられたコンテ

ンツの複製物に埋め込まれていたであろう識別情報を、より感度の高い符号によって推定可能とした電子透かし埋め込み装置、電子透かし解析装置、電子透かし埋め込み方法、電子透かし解析方法及びプログラムを提供することを目的とする。

【0024】

【課題を解決するための手段】

本発明に係る電子透かし埋め込み装置によれば、デジタルコンテンツの複製物に電子透かしとして埋め込むべき識別情報に一意に対応する、序列を持った複数のシンボルの並びからなるシンボル列を生成する第1の生成手段と、生成された前記シンボル列を構成する個々の序列のシンボルごとに、対応する埋め込み符号をそれぞれ生成する第2の生成手段と、生成された前記個々の序列のシンボルにそれぞれ対応する埋め込み符号を、前記デジタルコンテンツの複製物に電子透かしとして埋め込む符号埋め込み手段とを備えたことを特徴とする。

【0025】

また、本発明は、互いに異なる識別情報を電子透かしとして埋め込まれた複数のデジタルコンテンツの複製物をもとにして結託攻撃によって作出された不正のデジタルコンテンツの複製物を対象として、結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定するための電子透かし解析装置であって、前記デジタルコンテンツの複製物から、序列を持った複数の埋め込み符号をそれぞれ抽出する第1の抽出手段と、抽出された前記複数の埋め込み符号にそれぞれ対応するシンボルを求め、求められた複数のシンボル列を、対応する埋め込み符号の序列に従って並べてシンボル列を求める第2の抽出手段と、前記結託攻撃に用いられたデジタルコンテンツの複製物から抽出された前記シンボル列及び各々の正当な識別情報に対してそれぞれ一意に割り当てられるシンボル列に基づいて、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定する追跡手段とを備えたことを特徴とする。

【0026】

また、本発明は、デジタルコンテンツの複製物に電子透かしとして埋め込むべき識別情報に一意に対応する、序列を持った複数のシンボルの並びからなるシン

ボル列を生成し、生成された前記シンボル列を構成する個々の序列のシンボルごとに、対応する埋め込み符号をそれぞれ生成し、生成された前記個々の序列のシンボルにそれぞれ対応する埋め込み符号を、前記デジタルコンテンツの複製物に電子透かしとして埋め込むことを特徴とする。

【0027】

また、本発明は、互いに異なる識別情報を電子透かしとして埋め込まれた複数のデジタルコンテンツの複製物をもとにして結託攻撃によって作出された不正のデジタルコンテンツの複製物を対象として、結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定するための電子透かし解析方法あって、前記デジタルコンテンツの複製物から、序列を持った複数の埋め込み符号をそれぞれ抽出し、抽出された前記複数の埋め込み符号にそれぞれ対応するシンボルを求め、求められた複数のシンボル列を、対応する埋め込み符号の序列に従って並べてシンボル列を求め、前記結託攻撃に用いられたデジタルコンテンツの複製物から抽出された前記シンボル列及び各々の正当な識別情報に対してそれぞれ一意に割り当てられるシンボル列に基づいて、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定することを特徴とする。

【0028】

また、本発明は、コンピュータを電子透かし埋め込み装置として機能させるためのプログラムであって、デジタルコンテンツの複製物に電子透かしとして埋め込むべき識別情報に一意に対応する、序列を持った複数のシンボルの並びからなるシンボル列を生成する機能と、生成された前記シンボル列を構成する個々の序列のシンボルごとに、対応する埋め込み符号をそれぞれ生成する機能と、生成された前記個々の序列のシンボルにそれぞれ対応する埋め込み符号を、前記デジタルコンテンツの複製物に電子透かしとして埋め込む機能とをコンピュータに実現させるためのプログラムである。

【0029】

また、本発明は、互いに異なる識別情報を電子透かしとして埋め込まれた複数のデジタルコンテンツの複製物をもとにして結託攻撃によって作出された不正の

デジタルコンテンツの複製物を対象として、結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定するための電子透かし解析装置として、コンピュータを機能させるためのプログラムであって、前記デジタルコンテンツの複製物から、序列を持った複数の埋め込み符号をそれぞれ抽出する機能と、抽出された前記複数の埋め込み符号にそれぞれ対応するシンボルを求め、求められた複数のシンボル列を、対応する埋め込み符号の序列に従って並べてシンボル列を求める機能と、前記結託攻撃に用いられたデジタルコンテンツの複製物から抽出された前記シンボル列及び各々の正当な識別情報に対してそれぞれ一意に割り当てられるシンボル列に基づいて、前記結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定する機能機能とをコンピュータに実現させるためのプログラムである。

【0030】

なお、装置に係る本発明は方法に係る発明としても成立し、方法に係る本発明は装置に係る発明としても成立する。

また、装置または方法に係る本発明は、コンピュータに当該発明に相当する手順を実行させるための（あるいはコンピュータを当該発明に相当する手段として機能させるための、あるいはコンピュータに当該発明に相当する機能を実現させるための）プログラムとしても成立し、該プログラムを記録したコンピュータ読み取り可能な記録媒体としても成立する。

【0031】

本発明によれば、結託攻撃に用いられたコンテンツの複製物に埋め込まれていたであろう識別情報を、より感度の高い符号によって推定することができるようになる。また、符号長の短縮にも有効である。

【0032】

【発明の実施の形態】

以下、図面を参照しながら発明の実施の形態を説明する。

【0033】

同一のデジタルコンテンツの複製物（例えば、静止画、動画、音声、音楽等）の各々に対して、当該複製物ごとに異なる識別情報を透かし情報として埋め込み

、追跡する場合について説明する。以下では、識別情報として、当該複製物に対応するユーザすなわち当該複製物を利用することになるユーザ（例えば、当該複製物を記録媒体又は通信媒体を媒介として譲渡し若しくは貸し渡し又は提供する対象となるユーザ）のユーザ識別子（ユーザID）を用いる場合を主として説明するが、ユーザIDに対して所定の変換を施したものを識別情報としてもよいし、ユーザID以外の情報又はユーザID以外の情報を変換した情報を識別情報としてもよい。なお、ユーザIDには、利用の日時や利用場所等の情報を含めても良い。

【0034】

もちろん、デジタルコンテンツの複製物に対して、さらに、その他の様々な透かし情報（例えば、コンテンツの著作権者に関する情報、著作権等の権利に関する情報、コンテンツの利用条件に関する情報、その利用時に必要な秘密情報、コピー制御情報等、あるいはそれらを組み合わせたものなど）が様々な目的（例えば、利用制御、コピー制御を含む著作権保護、二次利用の促進等）で埋め込まれ、検出されてもよいが、その他の透かし情報を利用する場合における当該その他の透かし情報に関係する部分の構成は任意である。

【0035】

以下で示す構成図は、装置の機能ブロック図としても成立し、また、ソフトウェア（プログラム）の機能モジュール図あるいは手順図としても成立するものである。

【0036】

図1に、本発明の実施の形態に係る電子透かし埋込装置と電子透かし解析装置が適用されるシステムの概念図を示す。

【0037】

電子透かし埋込装置1と電子透かし解析装置2は、例えば、コンテンツ提供側に備えられ、管理される。

また、例えば、電子透かし埋込装置1は、ユーザ側に備えられ（例えば、コンテンツを利用するためのユーザ・システム（例えば、計算機システムや専用機器等に接続されあるいは組み込まれるなど）、電子透かし解析装置2は、コンテン

ツ提供側に備えられるようにしてもよい。

デジタルコンテンツの複製物への透かし情報の埋め込みは、前者の場合、該複製物がユーザへ渡されるのに先だって行われ、後者の場合、該複製物がユーザの利用に供されるのに先だって行われることになる。

【0038】

電子透かし埋込装置1においてデジタルコンテンツに所望の透かしデータを埋め込む方法や、電子透かし解析装置2においてデジタルコンテンツから該透かしデータ自体を取り出す方法は、どのような方法であってもよい（例えば、“松井甲子雄著、「電子透かしの基礎」、森北出版、1998年”等参照）。

電子透かし埋込装置1は、ソフトウェア（プログラム）としてもハードウェアとしても実現可能である。同様に、電子透かし解析装置2は、ソフトウェア（プログラム）としてもハードウェアとしても実現可能である。また、電子透かし埋込装置1および電子透かし解析装置2をコンテンツ提供側で用いる場合には、それらを一体化して実現することも可能である。

【0039】

図2に、電子透かし埋込装置1の構成例を示す。

【0040】

図2に示されるように、電子透かし埋込装置1は、埋め込むべき識別情報（例えば、ユーザID）に対応するシンボル列を生成するシンボル列生成部11と、シンボル列の各シンボルにそれぞれ対応する埋め込み符号（例えば、乱数列）を生成する埋め込み符号生成部12と、生成された埋め込み符号を対象コンテンツに埋め込む符号埋込部13とから構成される。

【0041】

電子透かし埋込装置1は、埋込対象コンテンツと、これに埋め込むべき識別情報（例えば、当該コンテンツを提供すべき対象ユーザのユーザID）とが与えられると、該識別情報に対応するシンボル列を生成し、生成されたシンボル列の各シンボルにそれぞれ対応する埋め込み符号を生成し、各埋め込み符号を対象コンテンツに埋め込んで、これを該識別情報に対応する複製物（例えば、該ユーザIDのユーザ向けの複製物）として出力する。他の透かし情報を利用する場合には

、必要に応じて他の透かし情報も埋め込まれる。

【0042】

なお、例えば、識別情報≠ユーザIDであって、かつ、ユーザIDが与えられるような場合には、与えられたユーザIDを識別情報に変換する、などの前処理を行う。

【0043】

電子透かし埋込装置1により得られた当該識別情報に対応するコンテンツの複製物は、記憶媒体や通信媒体などを媒介とした流通経路3を経てそれぞれ流通する。複数の複製物を用いた結託攻撃は、この流通経路3にて行われる。

【0044】

ここでは、結託攻撃としては、例えば、複数の複製物の単純平均操作（結託させる全コンテンツの重みを同一にして平均する）あるいは重み付け平均操作（結託させる全コンテンツの重みが全て同一ではないようにして平均する）などを想定している。

【0045】

図3に、電子透かし解析装置2の構成例を示す。

【0046】

図3に示されるように、電子透かし解析装置2は、埋め込み符号抽出部21と、シンボル列抽出部22と、追跡部23とを備えている。

【0047】

埋め込み符号抽出部21は、追跡対象コンテンツから、抽出すべきシンボル列の各シンボルにそれぞれ対応する埋め込み符号を抽出する。

【0048】

シンボル列抽出部22は、抽出された各抽出埋め込み符号からそれぞれ対応する抽出シンボルを求め、それら各抽出シンボルからなるシンボル列を出力する。

【0049】

追跡部23は、抽出シンボル列に対して、所定の追跡アルゴリズムを適用することによって、結託攻撃がなかったこと、または結託攻撃に用いられた複数の複製物にそれぞれ埋め込まれていた全識別情報（例えば、ユーザID）のうちの全

部又は一部を、特定（あるいは推測）しようとする。

【0050】

なお、例えば、識別情報≠ユーザIDであって、ユーザIDを求める必要があるような場合には、特定された識別情報を、さらに、ユーザIDに変換する、などの後処理を行う。

【0051】

なお、追跡アルゴリズム及びその前提となるシンボル列生成方法には種々のものがあり、例えば、一定の誤検出を許容して結託者の識別情報を求める確率論的方法や、誤検出なしに結託者の識別情報を求める決定論的方法などがある。

【0052】

以下、本実施形態についてより詳細に説明する。

【0053】

なお、以下では、識別情報=ユーザIDの場合を具体例にとって説明する（したがって、以下において、識別情報=ユーザIDとして説明している場合には、ユーザIDは一般的には識別情報のことを意味している）。

【0054】

また、以下では、各識別情報には、当該識別情報に固有となるシンボル列を割り当てるものとして説明する。なお、あるシンボル列に複数の識別情報が対応することを許す方法も可能である。

【0055】

まず、電子透かし埋込装置1について説明する。

【0056】

図4に、概略的な手順の一例を示す。

【0057】

シンボル列生成部11は、まず、対象複製物に埋め込むべき識別情報（本具体例では、ユーザID）に対応する、M（Mは複数）個のシンボルの並びからなるシンボル列{S(1), S(2), …, S(M)}を求める（ステップS1）。

【0058】

ここで、S(1)は、該識別情報に割り当てられたシンボル列すなわちシンボ

ルの並びにおける第1番目の序列に位置するシンボルすなわち第1シンボル（の値）を表し、 $S(2) \sim S(M)$ についても同様である。

【0059】

なお、該シンボル列は、各識別情報に対応して予め求めて記憶しておく方法と、必要時に生成する方法とがある。

【0060】

識別情報に対してシンボル列は一意に対応するので、ある正しい識別情報に対応するシンボル列 $\{S(1), S(2), \dots, S(M)\}$ が与えられると、該シンボル列に対応する識別情報が一意に特定されることになる。なお、結託攻撃を施された複製物から抽出した抽出シンボル列から、その結託攻撃に使用された識別情報を求めるのが、追跡アルゴリズムである。

【0061】

ここで、シンボルは、互いに区別できるものであれば、どのようなものでもよい。例えば、1つのシンボルとして、整数（または整数列）を用いてもよいし、アルファベット（またはアルファベット列）を用いてもよいし、英数字（または英数字列）を用いてもよいし、他のものでもよい。

【0062】

また、第1シンボル $S(1)$ の取り得る値（すなわち第1シンボル $S(1)$ の元）と、第2シンボル $S(2)$ の取り得る値と、…、第Mシンボル $S(M)$ の取り得る値とは、互いに異なるものであってもよいし、全て同一でもよい。例えば、シンボルとして整数を用いる場合に、第1シンボル $S(1)$ は $0 \sim 2$ のいずれかを取るものであり、第2シンボル $S(2)$ は $0 \sim 4$ のいずれかを取るものであり、第3シンボル $S(3)$ は $0 \sim 6$ のいずれかを取るものであり、…というように、互いに取り得る値の範囲の異なるものであってもよいし、第1シンボル $S(1)$ と第2シンボル $S(2)$ は $1 \sim 3$ のいずれかを取るものであり、第3シンボル $S(3)$ は $0 \sim 6$ のいずれかを取るものであり、…というように、一部は同じで一部は異なるようなものであってもよいし、 $S(1), S(2), \dots, S(M)$ の全てが $0 \sim 14$ のいずれかを取るものというように、全て同一でもよい。また、第1シンボル $S(1)$ には整数を用い、第2シンボル $S(2)$ にはアルファ

ベットを用い…、というような構成も可能である。

【0063】

以下では、具体例を用いて説明する場合には、シンボルに整数を用いる例を中心説明する。なお、例えば、まず整数を求め、求められた整数を対応するアルファベットに変換してそれをシンボルとして用いるものとする場合には、以下の説明において整数とアルファベットとの間の変換処理を追加すればよい。

【0064】

さて、与えられた識別情報から、それに一意に対応するシンボル列 {S (1), S (2), …, S (M)} を求める方法には、種々のバリエーションがある。

【0065】

例えば、第*i*のシンボルS (*i*) が、0～N (*i*) - 1の範囲内の整数値のいずれかを取るようにしてよい。ここで、N (1)、N (2)、…、N (M) は、例えば、予め定められた相互に異なる正整数である（互いに素の関係にある整数とするのが望ましい）。なお、N (1) < N (2) < … < N (M) としてもよいし、また、N (1) = N (2) = … = N (M) とする方法もあるし、他の方法もある。

【0066】

識別情報に対応するシンボル列の第*i*シンボルS (*i*) (*i* = 1～M) の各々には、0～N (*i*) - 1の範囲でランダムに値を割り当てる方法と、0～N (*i*) - 1の範囲で一定の規則に従って値を割り当てる方法とがある。また、いずれの場合においても、各識別情報には、互いにS (1)、S (2)、…、S (M) のうちの少なくとも一つが相違するように、M個のシンボルの組を排他的に割り当てるものとする。

【0067】

各識別情報に固有のシンボル列を割り当てるシンボル列生成方法には、例えば、識別情報の値として、0～N (1) × N (2) × … × N (M) - 1の範囲の整数の全部または一部を使用するものとし、対象となる識別情報UをN (*i*) で割ったときの剩余U mod N (*i*) を、該識別情報に対応するS (*i*) の値とする方法がある。この場合、これらM個の数N (1)、N (2)、…、N (M)

は、互いに素の関係にある整数とするのが望ましい（なお、以下の説明では、 $N(1) < N(2) < \dots < N(M)$ とする）。この場合に、詳しくは後述するよう、想定する最大の結託者数 c と、シンボル列の要素数 M と、使用する識別情報の範囲（例えば、0から始まる連番）との間に、一定の制約を与えるようにしてよい。

【0068】

ここで、この $S(i) = U \bmod N(i)$ によるシンボル列生成方法について数値を小さくとって簡単にした例を用いて説明する。

【0069】

例えば、シンボル列の要素の個数 M を 3 とし、 $N(1) = 3$, $N(2) = 5$, $N(3) = 7$ とする。この場合、第1シンボル $S(1)$ は 0~2 のいずれか、第2シンボル $S(2)$ は 0~4 のいずれか、第3シンボル $S(3)$ は 0~6 のいずれかとなる。

【0070】

次に、 $N(1) \times N(2) \times N(3) - 1 = 104$ であるので、0~104 の範囲の全部または一部をユーザIDとして用いる。ここでは、そのうち 0~14 をユーザIDとして用いるものとする。

【0071】

例えば、ユーザID = 7 の場合、

$$S(1) = 7 \bmod N(1) = 7 \bmod 3 = 1,$$

$$S(2) = 7 \bmod N(2) = 7 \bmod 5 = 2,$$

$$S(3) = 7 \bmod N(3) = 7 \bmod 7 = 0,$$

となる。

【0072】

また、例えば、ユーザID = 8 の場合、

$$S(1) = 8 \bmod N(1) = 8 \bmod 3 = 2,$$

$$S(2) = 8 \bmod N(2) = 8 \bmod 5 = 3,$$

$$S(3) = 8 \bmod N(3) = 8 \bmod 7 = 1,$$

となる。

【0073】

図5 (a) に、この例において各識別情報（ユーザID=0～14）について求められたシンボル列の各序列のシンボル $S(1)$, $S(2)$, $S(3)$ を示す。

【0074】

なお、シンボルにアルファベットを用いた場合には、例えば、図5 (b) のようになる（なお、図5 (b) では、 $S(1)=0$ にも $S(2)=0$ にも $S(3)=0$ にも全て a を対応させているが、例えば、 $S(1)=0$ には a を対応させ、 $S(2)=0$ には b を対応させ、 $S(3)=0$ には c を対応させるような構成も可能である）。

【0075】

図5 (a) において、例えば、ユーザID=7の場合、 $S(1)=1$ 、 $S(2)=2$ 、 $S(3)=0$ であるから、ユーザID=7に対応するシンボル列は、

{1, 2, 0}

となる。なお、シンボルにアルファベットを用いる場合には、例えば、

{b, c, a}

となる。

【0076】

また、例えば、ユーザID=8の場合、 $S(1)=2$ 、 $S(2)=3$ 、 $S(3)=1$ であるから、ユーザID=7に対応するシンボル列は、

{2, 3, 1}

となる。なお、シンボルにアルファベットを用いる場合には、例えば、

{c, d, b}

となる。

【0077】

図6に、図5の各識別情報（ユーザID=0～14）に対応するシンボル列を示す（整数を用いた場合と、アルファベットを用いた場合の両方を示す）。

【0078】

さて、次に、埋め込み符号生成部12は、シンボル列生成部11により生成さ

れた、対象複製物に埋め込むべき識別情報（本具体例では、ユーザID）に対応するシンボル列（S（1）、S（2）、…、S（M））を受けて、第1シンボルS（1）に対応する第1埋め込み符号R（1）、第2シンボルS（2）に対応する第2埋め込み符号R（2）、…、第MシンボルS（M）に対応する第M埋め込み符号R（M）をそれぞれ生成する（ステップS2）。各識別情報に対応するシンボル列の各シンボルに対応する埋め込み符号は、予め生成して記憶しておく方法と、必要時に生成する方法とがある。

【0079】

ここで、第1シンボルS（1）において、第1シンボルS（1）の値として取り得る各々の元*i*に対応する第1埋め込み符号 $w_1(i)$ は、それぞれ互いに、相互相関が無いようにする（ $w_1(i)$ は、第1シンボルにおいて、あるシンボルの元*i*に対応する埋め込み符号を表す）。つまり、 $i \neq j$ ならば $w_1(i) \cdot w_1(j) = 0$ 、 $i = j$ ならば $w_1(i) \cdot w_1(i) = 1$ とする。例えば、第1シンボルS（1）の取り得る値が{a}または{b}または{c}である場合に、第1シンボルS（1）= {a}に対応する埋め込み符号と、第1シンボルS（1）= {b}に対応する埋め込み符号と、第1シンボルS（1）= {c}に対応する埋め込み符号とは、いずれの間にも、相互相関が無いようにする。第kシンボル（ $k = 2 \sim M$ ）についてもそれぞれ同様である（ $i \neq j$ ならば $w_k(i) \cdot w_k(j) = 0$ 、 $i = j$ ならば $w_k(i) \cdot w_k(i) = 1$ とする）。なお、相互相関が無いような符号を用いる代わりに、相互相関が非常に小さくなるような符号を用いるようにしてもよい（後者の方が実用的である）。

【0080】

なお、識別情報に対応するシンボル列を構成する第1～第Mシンボルに対応する第1～第M埋め込み符号をそれぞれコンテンツに埋め込むにあたって、それら第1～第M埋め込み符号を互いに異なる箇所に埋め込むなど、互いに影響あるいは干渉を及ぼさない形態をとる場合には、互いに他の序列のシンボルに係る埋め込み符号との相互間、すなわち、第1シンボルに用いる埋め込み符号と、第2シンボルに用いる埋め込み符号と、…、第Mシンボルに用いる埋め込み符号との相互間には、特に上記のような相互相関についての制約はない。

【0081】

これに対して、識別情報に対応するシンボル列を構成する第1～第Mシンボルに対応する第1～第M埋め込み符号の全部又は一部の複数のものを同一の箇所に重畠するなど、通常であれば互いに影響あるいは干渉を及ぼし得る形態をとる場合には、それら埋め込み符号については、相互相関が無いようにするか、あるいは相互相関が非常に小さくなるようにするものとする。例えば、第1シンボルに対応する第1埋め込み符号～第Mシンボルに対応する第M埋め込み符号の全てを同一の箇所に重畠する場合には、第1シンボルの全元に対応する埋め込み符号～第Mシンボルの全元に対応する埋め込み符号の全ての埋め込み符号の相互間について、相互相関が無いようにするか、あるいは相互相関が非常に小さくなるようにする。

【0082】

この埋め込み符号としては、種々のものを用いることができる。

【0083】

例えば、特開2001-285623号公報に記載されたシンプレクティック符号を用いることができる。シンプレックス符号とは、符号長 $n - 1$ 、符号語数 n で、符号語間の相互相関が $-1/(n - 1)$ となる符号であり、 n 次のアダマール(Hadamard)行列を基に構成することができる。すなわち、符号語が $n - 1$ 次元ユークリッド空間中の $n - 1$ 次元単体の頂点に位置するような符号がシンプレックス符号である。例えば、3次元ユークリッド空間の場合は、図7に示すように $(-1, -1, 1)$, $(1, -1, -1)$, $(-1, 1, -1)$ で示す3つの頂点に位置する符号がシンプレックス符号を構成する。

【0084】

さて、このようにして生成された識別情報に対応するシンボル列の各シンボルにそれぞれ対応する第1～第M埋め込み符号は、電子透かし埋込装置1の符号埋込部13が、当該埋め込み符号の内容に従ってコンテンツを予め定められた方法で微少に変更することによって、対象コンテンツに埋め込まれる（ステップS3）。なお、コンテンツの変更の仕方には、まず埋め込み対象コンテンツに対して第1埋め込み符号R(1)を埋め込むようにコンテンツを変更し、次いで、該第

1埋め込み符号R(1)を埋め込まれたコンテンツに対して第2埋め込み符号R(2)を埋め込むようにコンテンツを変更し…、というように、逐次的に埋め込みを行う方法や、埋め込み対象コンテンツに対して一挙に全埋め込み符号R(1)～R(M)を埋め込むようにコンテンツを変更するようにする方法や、その他の方法がある。

【0085】

前述したように、生成された埋め込み符号を対象コンテンツに埋め込む方法（コンテンツのどのような量をどのように変更するか）には、特に限定はなく、どのような方法でも本発明は適用可能である。

【0086】

ただし、（識別情報に対応するシンボル列の）各シンボルにそれぞれ対応する埋め込み符号は、互いに影響が無いあるいは小さいように、符号構成や埋め込む場所や埋め込み方法等が選択されているようにする。

【0087】

次に、電子透かし解析装置2について説明する。

【0088】

ここで、結託攻撃及び結託者の追跡について、数値を小さくとて簡単にした例を用いて説明する。

【0089】

本実施形態では、結託攻撃として例えば複数の複製物の単純平均操作あるいは重み付け平均操作などを想定している。

【0090】

例えば、図6のユーザID=0, 3, 10の各ユーザが入手したコンテンツには、それぞれ、シンボル列として、

{a, a, a}

{a, d, d}

{b, a, d}

が埋め込まれている。

【0091】

ここで、上記の3人によって3つのコンテンツ複製物の単純平均操作による結託攻撃が行われたとする。

【0092】

本実施形態では、識別情報に対応するシンボル列の第1シンボル～第Mシンボルのそれぞれにおいて、異なる元に対応する埋め込み符号を互いに相關の無いあるいは非常に小さい電子透かしとして埋め込まれている。しかし、抽出すべきシンボル列における第1シンボル～第Mシンボルのそれぞれについて、最大剩余及び最小剩余（＝剩余対）ではなく、埋め込み符号（すなわち、第1埋め込み符号～第M埋め込み符号）を検出し、埋め込み符号をシンボルに戻して、シンボル列を抽出する。ここで、シンボル列における各序列のシンボルをそれぞれ検出するにあたって、結託攻撃に実際に使用されたコンテンツに埋め込まれていた複数のシンボルの元が検出されることがある（上記の具体例では、第1シンボルについて、aとb、第2シンボルについて、aとd、第3シンボルについて、aとdが検出され得る）。本実施形態では、第1～第Mのシンボルの各々について、結託攻撃に最も数多く用いられたシンボルがいずれかを評価し、最大人口（highest population）と評価されるシンボルを検出し、その検出された最大人口シンボル（highestly populated alphabet）によるシンボル列に基づいて、結託者の識別情報を追跡する。

【0093】

なお、抽出すべきシンボル列のある序列のシンボルにおいて、最大人口シンボルとなる元が複数存在する場合には、いずれか1つのみを最大人口シンボルとして検出するようにしてもよいし、それら複数を最大人口シンボル（群）として検出するようにしてもよい。

【0094】

また、最大人口シンボル（群）に加えて、最大人口シンボル以外で検出されたシンボルをも考慮して、追跡を行うようにしてもよい。

【0095】

上記の具体例の場合、第1～第3のシンボルのそれぞれについて、最も結託に用いられた数の多いシンボルの元が各シンボルの値として検出されるようなシン

ボル列が埋め込まれたコンテンツが作出されることになり、例えば、上記の例では、

{ a, a, d }

というシンボル列が埋め込まれたコンテンツ複製物が作出される。この場合、本実施形態の電子透かし解析装置2は、{ a, a, d } というシンボル列を検出する。

【0096】

ここで、図6を見てみると、{ a, a, d } というシンボル列に対応する識別情報は存在しない。従って、正当な識別情報が改変されていることがわかる。また、最大人口シンボル以外にも、第1シンボルでは元b、第2シンボルでは元d、第3シンボルでは元aに対応する埋め込み符号を検出することができるので、このことからも、結託攻撃を受けたことがわかる。

【0097】

しかし、本実施形態の電子透かし解析装置2は、検出された最大人口シンボルによるシンボル列 { a, a, d } (または、これに加えて、最大人口シンボル以外で検出されたシンボル)に基づいて、結託者の識別情報を追跡することができる。

【0098】

上記した例の場合、追跡アルゴリズム及びその前提となるシンボル列生成方法の構成方法に応じて、例えば、ある誤認識率を含んだ上で、例えばユーザID=0のみ(あるいは例えばID=0, 3, 10の全て)が特定され、あるいは誤認識なしに例えばユーザID=0のみ(あるいは例えばID=0, 3, 10の全て)が特定されることも可能になる。

【0099】

さて、追跡アルゴリズムとしては、様々なバリエーションが可能であるが、基本的には、抽出されたシンボル列に基づいて、結託攻撃に使用された複製物に埋め込まれていたであろう埋め込み符号／シンボル列に対応するユーザIDを求め、これを結託攻撃を行った結託者のユーザIDとして特定する。

【0100】

本実施形態では、結託攻撃後のコンテンツにおいて検出される最大人口によるシンボル列の第 i シンボルの値は、結託攻撃を行った複数のユーザのうちのいずれかのユーザのユーザIDに対応するシンボル列の第 i シンボルの値に一致する。よって、対象コンテンツから検出された最大人口による埋め込み符号から得られたシンボル列を解析することによって、かりにすべての結託者のユーザIDが特定できなかったとしても、一部の結託者のユーザIDを特定できることが期待される。

【0101】

また、抽出されたシンボル列が、いずれかの正当なユーザに対応するシンボル列と完全に一致する場合には、結託攻撃はなされていないことを特定することができる。ただし、抽出された最大人口によるシンボル列がいずれかの正当なユーザに対応するシンボル列と完全に一致しても、最大人口シンボル以外にも他のシンボルに対応する埋め込み符号が検出されている場合には、結託攻撃を行ったがシンボル列の改変に失敗したことが考えられ、この場合には、特定されたユーザIDは、結局、結託攻撃を行っていたことになる。

【0102】

追跡アルゴリズムの代表的な種類としては、確率論的方法や決定論的方法などがある。

【0103】

確率論的方法は、基本的には、対象コンテンツから検出されたシンボル列から選択された所定数のシンボルが、あるユーザIDを示していれば（すなわち、検出シンボル列から選択された所定数のシンボルの値と、あるユーザIDに対応する識別情報に対応するシンボル列の同じ序列位置のシンボルの値とがそれぞれ一致すれば）、それを結託者のユーザIDとして出力するものである。上記の所定数は、例えば、誤認識率等に応じて決定される（なお、誤認識は、例えば、該選択された所定数のシンボルが実際には一人のユーザにのみ属するものではなく、複数のユーザに分属するものである場合に、結託者でないユーザのユーザIDが偶然に出力されてしまうことで生じる）。誤認識をより低く抑えるためには、例えば、上記の所定数をより大きな値に設定すればよい。

【0104】

決定論的方法は、基本的には、検出されたシンボル列から、誤認識なしに結託者のユーザIDを特定できるようにするものである。すなわち、検出されたシンボル列が結託攻撃によって作出されるために必要不可欠であった結託者のユーザIDを解析的に特定するとともに、特定できない場合には追跡不能とする方法である。追跡不能（結託者のユーザIDがすべて特定不能）となるケースを0にするまたはより低く抑えるためには、例えば、ユーザIDの取り得る値の範囲を一定とした場合にシンボル列の要素数をより多くすればよい。

【0105】

図8に、概略的な手順の一例を示す。

【0106】

本実施形態の電子透かし解析装置2において、まず、埋め込み符号抽出部21は、検出対象となるコンテンツの複製物から埋め込み符号R' (1)、R' (2)、…、R' (M) を抽出する（ステップS21）。そして、シンボル列抽出部22は、抽出された埋め込み符号R' (1)、R' (2)、…、R' (M) から、対応するシンボル列{S' (1)、S' (2)、…、S' (M)} を求める（ステップS22）。

【0107】

図9に、埋め込み符号抽出部21の処理手順の一例を示す。

【0108】

埋め込み符号抽出部21は、追跡対象となるコンテンツの複製物から抽出すべきシンボル列の各シンボルS' (1)、S' (2)、…、S' (M) にそれぞれ対応する埋め込み符号R' (1)、R' (2)、…、R' (M) の各々について、当該埋め込み符号R' (i) として使用されている可能性のある各符号（すなわち、シンボルS' (i) の各元にそれぞれ対応する各埋め込み符号）の評価値をそれぞれ求め（評価値計算処理）（ステップS31）、各埋め込み符号R' (i) ごとに、求めた各符号の表価値に基づいて、第iの埋め込み符号を決定する（HPE (Highestly Populated Elements) 計算処理）（ステップS32）。

【0109】

ここで、平均による結託攻撃の場合、 $\Omega((L/\ln n)^{1/2})$ 人の結託によって電子透かしの消去が行えるという解析結果がある（例えば文献1「J.Kilian, F.T.Leighton, T.G.Shamoon, R.E.Tarjan and F.Zane, "Resistance of Digital Watermarks to Collusive Attacks" , NEC Research Institute, Technical Report TR-585-98, 1998.」）。ここで、Lはコンテンツのサイズ（電子透かしを埋め込める容量）、nは配布先となるユーザの総数とする。これより、最大結託サイズをcとすれば、長さL（Lは $c^2 \cdot \ln n$ に比例する値）のコンテンツが必要となるため、cが大きくなると電子透かしが失われる。この解析は、電子透かしがランダムなガウシアンノイズとして重畠されることを仮定しており、平均による結託攻撃により電子透かしの平均は、中心極限定理によりcが大きくなると0になる。

【0110】

そこで、本実施形態では、前述したように、各ユーザIDに対応するシンボル列における第*i*シンボル（*i*=0～M）に対して、N(*i*)個の元の集合= { $a_{i,0}, a_{i,1}, \dots, a_{i,N(i)-1}$ } のうちから選択した1つの元 $a_{i,j}$ に対応する埋め込み符号 $r_{i,j}$ （ $r_{i,j}$ は他の元 $a_{i,k}$ （ $k \neq j$ ）に対応する埋め込み符号 $r_{i,k}$ （ $k \neq j$ ）との間に相関の無いあるいは非常に小さい埋め込み符号）を電子透かしとして埋め込み、平均による結託攻撃が行われた場合には、剩余対（=最大剩余及び最小剩余）の集合ではなく、最大人口による第*i*シンボルを検出し、最大人口シンボルからなる抽出シンボル列を検出する方法をとっている。

【0111】

例えば文献1の解析では、電子透かしが一般的なガウシアンノイズとしてモデルされているが、この方法では、そのようなガウシアンノイズの成す空間のうち、非常に限定された（例えば、N(*i*)次元の）部分空間の中にしか電子透かしが値を取らない。しかも、それぞれの成分が直交しているため、平均をとっても、複数の電子透かしが相殺しあうことはない。

【0112】

最大人口シンボルに対応する埋め込み符号の抽出は、例えば、次のように行うことができる。

【0113】

抽出すべきシンボル列における第*i*シンボルに対して、上記の評価値計算処理として、各元 $a_{i,j}$ （ここで、 $j = 0 \sim N(i-1)$ ）に対して相互相関の無い埋め込み符号 $m_{i,j}$ が対応している（つまり、 $m_{i,j} \cdot m_{i,k} = \delta_{j,k}$ とする）。第*i*シンボルについて、コンテンツ*I*と、第*i*シンボルにおける各元 $a_{i,j}$ に対応する埋め込み符号 $m_{i,j}$ との相互相関 $C_{i,j} = m_{i,j} \cdot I$ を測定する。この相関値 $C_{i,j}$ を、第*i*シンボルに対応する第*i*埋め込み符号における当該符号 $m_{i,j}$ の評価値とする。この値 $C_{i,j}$ が最大となる j_{max} に係る埋め込み符号 $m_{i,j_{max}}$ が、第*i*埋め込み符号における最大人口シンボルに対応する埋め込み符号であり、これに対応する第*i*シンボルの元 $a_{i,j_{max}}$ が、求めるべき第*i*シンボルの最大人口シンボルとなる。

【0114】

図10は、抽出すべきシンボル列のある序列のシンボル（第1シンボルとする）に係る最大人口シンボルの決定を説明するための図である。第1シンボルの各元（本例では、 a, b, c ）にそれぞれ対応する第1の埋め込み符号の候補としての各符号についてそれぞれ求められた評価値を、ベクトル v で表すものとする。図10において、各軸は、第1シンボルの元に対応する。ベクトル v の方向が最も近い軸に対応する元を、求めるべきシンボル列における第1シンボルとして出力する。

【0115】

例えば、図11(a)のように、単純平均化法による結託攻撃に使用されたコンテンツの複製物の数が100（あるいは、重み付け平均化法による重みを考慮した延べ数が100）であり、第1シンボルにおいて、元 $a_{1,1}, a_{1,2}, a_{1,3}, a_{1,4}$ に対応する人口（結託攻撃に使用された当該元に係るコンテンツの複製物の数）がそれぞれ、53, 23, 8, 13であったとすると、最大人口の第1シンボル $a_{1,1}$ に対応する埋め込み符号 $r_{1,1}$ が、埋め込み符号抽出部21によつて求められる。そして、シンボル列抽出部22は、抽出された第1の埋め込み符号 $r_{1,1}$ を受け、これに対応するシンボル $a_{1,1}$ を第1シンボルとする。同様にして、最大人口による第2～第Mシンボルが求められ、抽出すべきシンボル列が得

られる。なお、結託攻撃に使用されたコンテンツの複製物に埋め込まれている識別情報の組み合わせによっては、図11（b）のように、ほとんど1つのシンボルに偏ることも、唯一のシンボルに偏ることもあり得る。

【0116】

なお、前述したように、抽出すべきシンボル列のある序列のシンボルにおいて、最大人口シンボルとなる元が複数存在する場合には、いずれか1つのみを最大人口シンボルとして検出するようにしてもよいし、それら複数を最大人口シンボル（群）として検出するようにしてもよい。

【0117】

また、最大人口シンボル（群）に加えて、最大人口シンボル以外で検出されたシンボルをも考慮して、追跡を行うようにしてもよい。

【0118】

次いで、追跡部23は、抽出されたシンボル列{S'（1）、S'（2）、…、S'（M）}に対して、追跡アルゴリズムを実行する（ステップS23）。

【0119】

まず、確率論的方法による追跡アルゴリズムの一例について説明する。

【0120】

最大人口シンボル（群）からなるシンボル列に基づいて確率論的に結託者を特定する方法は、例えば、c-s e c u r e CRT符号において、マーキング仮定を変更することで得られる。ここでは、第*i*シンボルにおいて、最大人口シンボルとしてある元 $a_{i,j}$ が検出される確率は、 $1/N(i)$ であるという仮定をおく。これは、結託がランダムに構成されるとすれば妥当な仮定である（あるいは、第*i*シンボル内でシンボルの元の間でランダムな置換を行うこととし、検出時の確率は、すべてのランダムな置換にわたってとることとすれば、妥当な仮定となる）。

【0121】

図12に、この場合の処理手順の一例を示す。

【0122】

まず、すべてのユーザID（= U_i とする）について、 $D(U_i)$ を求める（ス

ステップS41)。

【0123】

ここで、 $D(U_i)$ は、該ユーザ ID = U_i のユーザに割り当てられるシンボル列 $\{S(1), S(2), \dots, S(M)\} = \{U_i \bmod N(1), U_i \bmod N(2), \dots, U_i \bmod N(M)\}$ と、抽出されたシンボル列 $\{S'(1), S'(2), \dots, S'(M)\}$ とを、両シンボル列で同一の序列に位置するもの同士でそれぞれ比較した場合に、一致したシンボルの個数を表す。

【0124】

次に、各ユーザ ID = U_i に対して求めた $D(U_i)$ について、当該 $D(U_i)$ と、あらかじめ定められたしきい値 D_{th} とを比較し、 $D(U_i) \geq D_{th}$ ならば、そのユーザ ID = U_i を、結託者のユーザ ID と特定する（ステップS42）。

【0125】

なお、 D_{th} は、例えば、 $D_{th} = k + 1$ で、 k は、 $N(0), \dots, N(k)$ の積を、ユーザ ID（識別情報）の総数以上とするような値（ここで、 $N(0) \leq N(1) \leq \dots \leq N(M)$ とする）であり、1は、例えば、式（1）で与えられる。

$$[1 - \prod_{i=1}^{k+1} \frac{1}{N(i)}]^S \geq 1 - \varepsilon_2 \quad (1)$$

ここで、 \prod をとる i の範囲は、 $i = 0 \sim (k+1-1)$ あるいは $i = k \sim (k+1-1)$ 、

$$S = M C_{k+1}$$

ε_2 は、結託者のユーザ ID（識別情報）における追跡誤り率で、 $0 < \varepsilon_2 < 1$ を満たす。

【0126】

例えば、前述の具体例において、図6のユーザ ID = 0, 3, 10 の各ユーザが入手したコンテンツには、それぞれ、シンボル列として、

{a, a, a}

{a, d, d}

{b, a, d}

が埋め込まれており、それら3人によって3つのコンテンツ複製物の単純平均操

作による結託攻撃が行われて、

{ a, a, d }

というシンボル列が埋め込まれたコンテンツ複製物が作出されたものとする。

【0127】

ここで、{ a, a, d } という抽出シンボル列が得られたとする。

【0128】

ユーザID=0について、これに固有に割り当てられたシンボル列と、抽出シンボルとを比較すると、第1シンボルはいずれも { a } で一致し、第2シンボルはいずれも { a } で一致し、一方、第3シンボルは { a } と { d } で一致しないので、この場合、 $D(0) = 2$ となる。同様に、 $D(3) = 2$ 、 $D(10) = 2$ となる。

【0129】

ここで、かりに $D_{th} = 2$ ならば、図6のユーザID=0～14のうち、ユーザID=0, 3, 10の3人のみが、 $D(0) \geq D_{th}$ 、 $D(3) \geq D_{th}$ 、 $D(10) \geq D_{th}$ を満たすので、この場合には、ユーザID=0, 3, 10が、結託者のユーザIDとして特定される。

【0130】

ここで、従来のc-secure CRT符号は、各内符号が複数のブロックで構成され、各ブロックのビット反転を観測することで、最大剩余及び最小剩余を検出しておらず、さらにブロックを構成する各ビットは、例えば、電子透かしとして埋め込まれることを想定していたのに対して、本実施形態では、シンボル列の元が電子透かしとして埋め込まれており、最大人口シンボルを検出することで、内符号感度の問題を軽減できるとともに、符号の構成が単純になったため符号長の短縮にもつながる。前述したように、外符号としては、c-secure CRT符号の外符号のマーキング仮定の変更を反映したc-secure性のための条件から決定された個数の内符号を用意することで、c-secure CRT符号と同様の追跡が行える。

【0131】

ここで、文献2「Y.Yacobi, "Improved Boneh-Shaw content fingerprinting

” , Topics in Cryptology - C-RSA 2001, 378-391, 2001.」に提案されている方法との差異を述べておく。文献2では、内符号のブロックを電子透かしで置き換える方法を提案している。この場合、最大剩余及び最小剩余を求める内符号感度の問題は解決されていない。しかし、文献2では、外符号として、文献3 「D. Boneh and J. Shaw, “Collusion-Secure Fingerprinting for Digital Data” , Advance in Cryptology: Proceedings of CRYPTO' 95, Springer-Verlag, 452-465, 1995.」や文献4 「D. Boneh and J. Shaw, “Collusion-Secure Fingerprinting for Digital Data” , IEEE Transactions on Information Theory, 44, 5, 1897-1905, 1998.」で提案されているような誤り訂正符号を用いた外符号、あるいは、ランダム符号による符号の構成を前提としている。これらの外符号では、最大剩余及び最小剩余は必要ではなく、剩余の中から1つ検出できていればよい。ただし、この外符号は、 $O(c^2 \cdot \log n)$ の長さを持つことから、cが大きくなると急速に符号長が大きくなる。

【0132】

次に、決定論的方法による追跡アルゴリズムの一例について説明する。

【0133】

決定論的方法による追跡アルゴリズムは、例えば、文献5 「H. Muratani, “Combination Outer Codes for c-secure CRT Codes” , SCIS2002, 13D-3, 2002.」の方法を利用することができる。この場合、想定する最大の結託者数cが大きいときは埋め込み符号の符号長をある程度符号長を長くするのが好ましいが、cが小さいときは確率論的方法による追跡アルゴリズムにおける埋め込み符号よりも短い符号長で済むことになる。

【0134】

誤検出を0または非常に小さくするために、想定する最大の結託者数cと、シンボル列の要素数Mと、使用する識別情報の範囲（例えば、0から始まる連番）との間に、一定の制約を与えるようにしてもよい。例えば、 $M > c^2 \cdot (K - 1)$ とするようにしてもよい。

【0135】

ここで、追跡アルゴリズムでは、結託攻撃が行われる際の最大結託数をcと想

定する。すなわち、 $2 \leqq a \leqq c$ を満たす任意の a 人の結託者すなわち a 個の相互に異なる識別情報の埋め込まれたコンテンツの複製物により結託攻撃が行われる場合を考える（なお、ここでは、ユーザと識別情報とが 1 対 1 に対応するものとして考えているが、例えば、ユーザと識別情報とが 1 対多対応となり得る場合であって、実際に同一の結託者が自身に対応する 2 つの識別情報をそれぞれ埋め込まれた同一のコンテンツについての 2 つの複製物を当該結託攻撃に供する際には、結託者の数とは、異なる識別情報を埋め込まれたコンテンツの複数物ごとに考えることになる）。

【0136】

以下では、 a ($2 \leqq a \leqq c$) 個の相互に異なる識別情報の埋め込まれたコンテンツの複製物により結託攻撃を行う際に供される当該複数のコンテンツの複製物にそれぞれ埋め込まれた識別情報により構成される識別情報群を、結託者グループと呼ぶものとする。例えば、図 6 の例では、識別情報 = ユーザ ID = 0 ~ 14 であるので、最大結託数 $c = 3$ とすると（実際には識別情報の範囲や最大結託数はもっと大きく取られる）、結託者グループには、{0, 1}、{0, 2}、…、{0, 14}、{1, 1}、{1, 2}、…、{1, 14}、{2, 3}、…、{2, 14}、{3, 4}、…、{11, 14}、{12, 13}、{12, 14}、{13, 14}（以上、 $a = 2$ による結託）、{0, 1, 2}、{0, 1, 3}、…、{0, 1, 14}、{0, 2, 3}、{0, 2, 4}、…、{0, 2, 14}、{0, 3, 4}、…、{10, 11, 14}、{11, 12, 13}、{11, 12, 14}、{12, 13, 14}（以上、 $a = 3$ による結託）がある。最大結託数 $c = 3$ 以下の結託攻撃は、必ず、上記の結託者グループのいずれかによって行われる。

【0137】

次に、追跡対象コンテンツから抽出された抽出シンボル列を入力し、これに適合する結託者グループを出力する。

【0138】

ここで、ある結託者グループが抽出シンボル列に適合するとは、次の場合である。すなわち、その結託者グループを構成する複数の識別情報の各々に固有に対

応するシンボル列を基にして、何らかの結託による操作（例えば平均操作）により、当該抽出シンボル列が作出され得るものと判断される場合に、当該ある結託者グループが、当該抽出シンボル列に適合すると判断する。

【0139】

例えば、図6の具体例において、あるコンテンツから、

{ a, a, d }

というシンボル列が抽出された場合に、前述の図6のユーザID = 0, 3, 10 からなる結託者グループは、当該抽出シンボル列 { a, a, d } に適合することになる。

【0140】

適合する結託者グループの数は、抽出シンボル列の内容によって、1つになる場合と、複数になる場合とがあり得る。

【0141】

なお、抽出されたシンボル列が、いずれかの正当なユーザに対応するシンボル列と完全に一致し、かつ、最大人口シンボル以外のシンボルが存在しない場合には、結託攻撃はなされていないことを特定することができる。ただし、抽出された最大人口によるシンボル列がいずれかの正当なユーザに対応するシンボル列と完全に一致しても、最大人口シンボル以外にも他のシンボルに対応する埋め込み符号が検出されている場合には、結託攻撃を行ったがシンボル列の改変に失敗したことが考えられ、この場合には、特定されたユーザIDは、結局、結託攻撃を行っていたことになる。

【0142】

また、何らかの理由により、適合する結託者グループが得られない場合は、エラーとなる。

【0143】

さて、次に、抽出シンボル列に対して、1つの結託者グループのみが得られた場合には、該結託者グループを構成する識別情報を全て、結託者の識別情報とすればよい。

【0144】

一方、2以上の結託者グループが得られた場合には、得られた全ての結託者グループに共通に存在する（1又は複数の）識別情報のみを、結託者の識別情報とすればよい。なお、この共通に存在する識別情報を、共通識別情報と呼ぶものとする。

【0145】

ここで、同一の剩余対表現から複数の結託者グループが得られた場合には、いずれの結託者グループが、実際に結託攻撃に参加したであろう結託者のユーザID（識別情報）に対応するものかは、（抽出シンボル列のみからは）わからない。しかし、（結託攻撃が最大結託者数以下の結託者数で行われたとすれば）、それらのうちのいずれか1つの結託者グループが、求めるべき結託者グループである。したがって、得られた複数の結託者グループの全てに共通に存在する識別情報があれば、少なくともその識別情報（すなわち共通識別情報）は、実際に結託攻撃に参加したであろう結託者の識別情報である。

【0146】

なお、2以上の結託者グループが得られた場合であっても、共通識別情報が1つもないときは、特定不能とする。なお、得られた複数の結託者グループに共通識別情報が存在しないケースが理論的には発生しないように構成した場合には、得られた複数の結託者グループに共通識別情報が存在しないケースは何らかのエラーによってのみ発生することになる。

【0147】

このように、実際に結託攻撃に参加したであろう結託者が、たとえ全ては分からなくても、確実に一人でも分かるということは、デジタルコンテンツの違法コピーに対する事前の抑制や著作権侵害が発生したときの事後の救済のために、非常に大きな意味がある。

【0148】

ところで、複数の結託者グループが得られたものの共通識別情報が存在しないケースは、できるだけ少なく抑えらるのが望ましい。そして、このようなケースは、何らかの手段によって、少なくすることができます。

【0149】

例えば、複数の結託者グループが得られたものの、共通識別情報が存在しないようなケースを探索し、そのような結託者グループを構成する識別情報の全部又は一部を使用しないようにすることも、一つの方法である。

【0150】

また、例えば、想定する最大の結託者数cを一定とすれば、定性的には、整数の個数M（=成分符号の個数）を増加させるほど、あるいは使用する識別情報の範囲を減少させるほど、上記のケースを少なくすることができる。

【0151】

図13に、この場合の処理手順の一例を示す。

【0152】

まず、与えられた抽出シンボル列に適合する結託者グループを全て求める（ステップS51）。

【0153】

求められた結託者グループの数が1である場合（ステップS52）、求められた結託者グループを構成する全てのユーザIDを結託者のユーザID（結託者ID）とし（ステップS53）、求められた結託者IDを出力する（ステップS56）。

【0154】

求められた結託者グループの数が2以上である場合（ステップS52）、求められた全ての結託者グループに共通に存在するユーザIDがあれば（ステップS54）、求められた全ての結託者グループに共通に存在する（1又は複数の）ユーザIDのみを結託者IDとし（ステップS55）、求められた結託者IDを出力する（ステップS56）。一方、求められた全ての結託者グループに共通に存在するユーザIDがなければ（ステップS54）、結託者IDが得られなかった旨（またはその旨及び結託者グループを特定する情報、またはその旨及び抽出シンボル列）を出力する（ステップS57）。

【0155】

なお、ステップS2で結託者グループが1つも得られない場合には、エラーとなる。

【0156】

なお、ユーザIDの総数が大きい場合には、全ての結託者グループに対して、この処理を行うことは、コストが大きいので、例えば、結託者グループのサイズに関する条件や、現実の結託攻撃がそれを破る確率は十分に小さいと考えられる条件などによって、限定された結託者グループに対してのみ、この処理を行うようにもよい。

【0157】

さて、符号の構成方法としては、誤り訂正符号をc-secure CRT符号の外符号に用いる方法もある。この場合にも、決定論的な追跡を行う方法と、確率論的な追跡を行う方法がある。決定論的な追跡を行う外符号については、文献3、文献4、文献6「B.Chor, A.Fiat and M.Naor, “Tracing Traitors”, Advances in Cryptology-CRYPTO’ 94, LNCS 839, 257-270, 1994.」、文献7「H. D.L.Hollmann, J.H.van Lint, J.-P.Linnartz and L.M.G Tolhuizen, “On codes with the identifiable parent property”, Journal of Combinatorial Theory, 82, 121-133, 1998.」、文献8「J.N.Staddon, D.R.Stinson and R.Wei, “Combinational Properties of frameproof and treacability code”, 2000.」、文献9「R.Safavi-Naini and Y.Wang, “Collusion Secure q-ary Fingerprinting for Perceptual Content”, Workshop on Security and Privacy in Digital Rights Management 2001, November, 2001.」に符号の提案がある。

【0158】

確率論的な追跡については、従来のc-secure CRT符号（例えば、文献10「H.Muratani, “Collusion Resilience of digital watermarking”, SCSI2000, C06, 2000.」や文献11「H.Muratani, “Collusion-Secure Fingerprinting Code Reduced by Chinese Remaindering and its Random-Error Resilience”, Information hiding, Proceedings of the 4th International Workshop, IH 2001, 303-315, 2001.」）の外符号の拡張を行うことで構成できる。

【0159】

以下、Reed-Solomon符号を使った例を説明する。

【0160】

ここで、 $N = c(k + 1)$ とする。Cをnarrow sense [N, k, N - k + 1]_q Reed-Solomon 符号とする。

【0161】

次の式(2)の条件が満たされるとき、この符号は、確率論的外符号とできる

$$[1 - 1/q^k] S \geq 1 - \varepsilon \quad (2)$$

ここで、 $S = M C_{k+1}$ 、

$q = N(0) = N(1) = \dots = N(M)$ 、

ε は、結託者のユーザID(識別情報)における追跡誤り率で、 $0 < \varepsilon < 1$ なる実数とする。

【0162】

この場合に、先に例示した確率論的方法による追跡アルゴリズムの一例を適用することも可能である。なお、この場合には、先に例示した確率論的方法による追跡アルゴリズムの一例で、しきい値 $D_{th} = k + 1$ において、1は、式(1)の代わりに、例えば、式(2)で与えられるものとしてもよい。

【0163】

もちろん、この場合に、先に例示した決定論的方法による追跡アルゴリズムの一例を適用することも可能である。

【0164】

この符号は、AG (Algebraic Geometry) 符号をベースにした構成も可能である。

【0165】

以下では、本実施形態のハードウェア構成、ソフトウェア構成について説明する。

【0166】

本実施形態の電子透かし解析装置は、ハードウェアとしても、ソフトウェア(コンピュータに所定の手段を実行させるための、あるいはコンピュータを所定の手段として機能させるための、あるいはコンピュータに所定の機能を実現させるための)プログラム)としても、実現可能である。また、電子透かし解析装置

をソフトウェアで実現する場合には、記録媒体によってプログラムを受け渡しすることも、通信媒体によってプログラムを受け渡しすることもできる。もちろん、それらは、電子透かし埋込装置についても同様である。

また、電子透かし埋込装置や電子透かし解析装置をハードウェアとして構成する場合、半導体装置として形成することができる。

また、本発明を適用した電子透かし解析装置を構成する場合、あるいは電子透かし解析プログラムを作成する場合に、同一構成（あるいは、共有もしくは使い回し可能な構成）を有するブロックもしくはモジュールがあつても、それらをすべて個別に作成することも可能であるが、同一構成（あるいは、共有もしくは使い回し可能な構成）を有するブロックもしくはモジュールについては1または適当数のみ用意しておいて、それをアルゴリズムの各部分で共有する（使い回す）ことも可能である。電子透かし埋込装置を構成する場合、あるいは電子透かし埋め込みプログラムを作成する場合も、同様である。また、電子透かし埋込装置および電子透かし解析装置を含むシステムを構成する場合、あるいは電子透かし埋め込みプログラムおよび電子透かし検出プログラムを含むシステムを作成する場合には、電子透かし埋込装置（あるいはプログラム）と電子透かし解析装置（あるいはプログラム）に渡って、同一構成（あるいは、共有もしくは使い回し可能な構成）を有するブロックもしくはモジュールについては1または適当数のみ用意しておいて、それをアルゴリズムの各部分で共有する（使い回す）ことも可能である。

【0167】

また、電子透かし埋込装置や電子透かし解析装置をソフトウェアで構成する場合には、マルチプロセッサを利用し、並列処理を行って、処理を高速化することも可能である。

【0168】

また、これまで説明してきた構成は、装置の一部としてだけでなく、それ自体が一つの装置として成立可能である。例えば、電子透かし解析装置の復号部22は、電子透かし解析装置を構成する一体不可分の構成部分としても、電子透かし解析装置を構成する部品あるいはモジュール等としも、独立した復号装置として

も成立する。

【0169】

なお、この発明の実施の形態で例示した構成は一例であって、それ以外の構成を排除する趣旨のものではなく、例示した構成の一部を他のもので置き換えたり、例示した構成の一部を省いたり、例示した構成に別の機能あるいは要素を付加したり、それらを組み合わせたりすることなどによって得られる別の構成も可能である。また、例示した構成と論理的に等価な別の構成、例示した構成と論理的に等価な部分を含む別の構成、例示した構成の要部と論理的に等価な別の構成なども可能である。また、例示した構成と同一もしくは類似の目的を達成する別の構成、例示した構成と同一もしくは類似の効果を奏する別の構成なども可能である。

また、この発明の実施の形態で例示した各種構成部分についての各種バリエーションは、適宜組み合わせて実施することが可能である。

また、この発明の実施の形態は、個別装置としての発明、関連を持つ2以上の装置についての発明、システム全体としての発明、個別装置内部の構成部分についての発明、またはそれらに対応する方法の発明等、種々の観点、段階、概念またはカテゴリに係る発明を包含・内在するものである。

従って、この発明の実施の形態に開示した内容からは、例示した構成に限定されることなく発明を抽出することができるものである。

【0170】

本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

【0171】

【発明の効果】

本発明によれば、結託攻撃に用いられたコンテンツの複製物に埋め込まれていたであろう識別情報を、より感度の高い符号によって推定することができるようになる。

【図面の簡単な説明】

【図1】

本発明の一実施形態に係る電子透かし埋込装置及び電子透かし解析装置を含む
コンテンツ流通システムの概略構成を示す図

【図2】

同実施形態に係る電子透かし埋込装置の構成例を示す図

【図3】

同実施形態に係る電子透かし解析装置の構成例を示す図

【図4】

同実施形態に係る電子透かし埋込装置の概略的な手順の一例を示すフローチャート

【図5】

同実施形態に係る識別情報とシンボル列の具体例について説明するための図

【図6】

同実施形態に係る識別情報とシンボル列の具体例について説明するための図

【図7】

シンプレックス符号について説明するための図

【図8】

同実施形態に係る電子透かし解析装置の概略的な手順の一例を示すフローチャート

【図9】

同実施形態に係る埋め込み符号抽出部の処理手順の一例を示すフローチャート

【図10】

同実施形態に係る最大人口シンボルについて説明するための図

【図11】

同実施形態に係る最大人口シンボルについて説明するための図

【図12】

同実施形態に係る追跡部の処理手順の一例を示すフローチャート

【図13】

同実施形態に係る追跡部の処理手順の他の例を示すフローチャート

【符号の説明】

1 …電子透かし埋込装置

2 …電子透かし解析装置

3 …流通経路

1 1 …シンボル列生成部

1 2 …埋め込み符号生成部

1 3 …符号埋込部

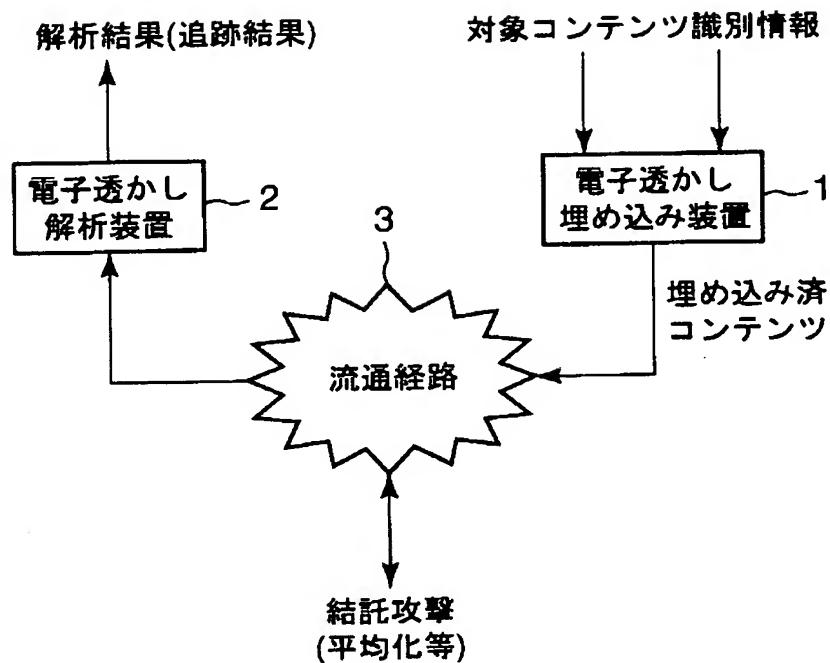
2 1 …埋め込み符号抽出部

2 2 …シンボル列抽出部

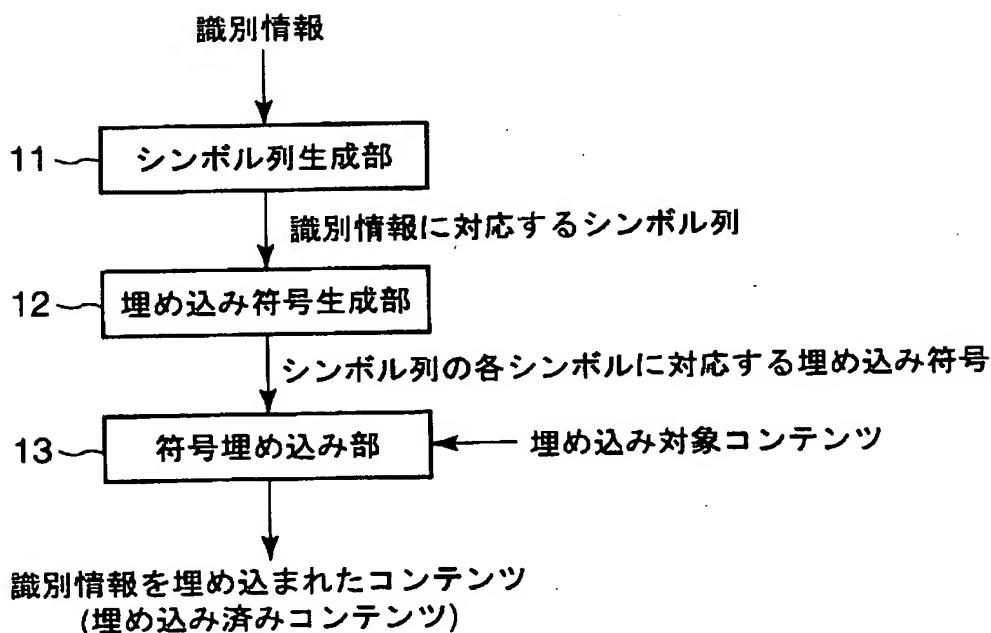
2 3 …追跡部

【書類名】 図面

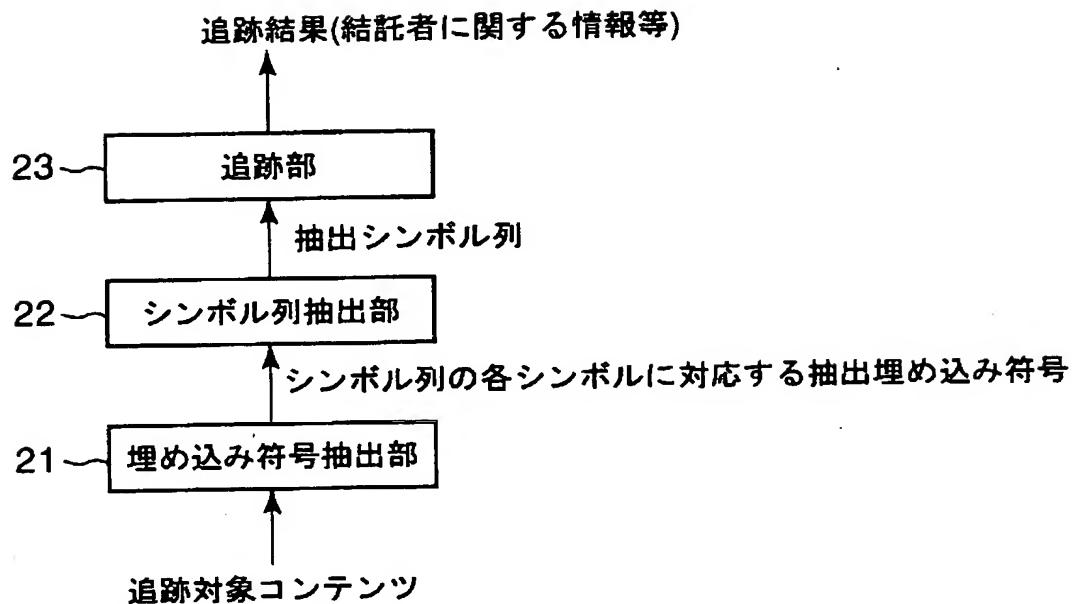
【図1】



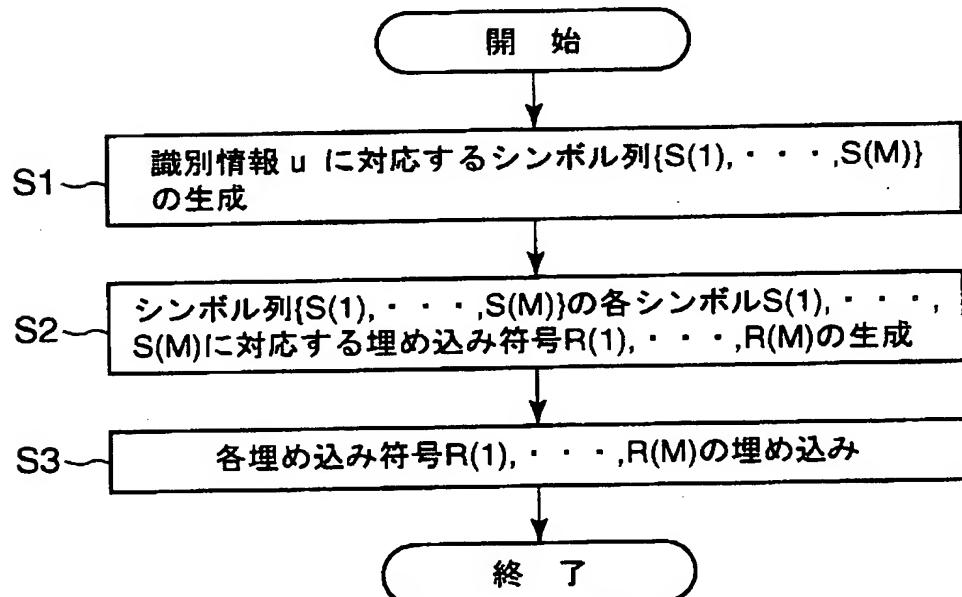
【図2】



【図3】



【図4】



【図5】

ID	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
S(1)	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
S(2)	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
S(3)	0	1	2	3	4	5	6	0	1	2	3	4	5	6	0

(a)

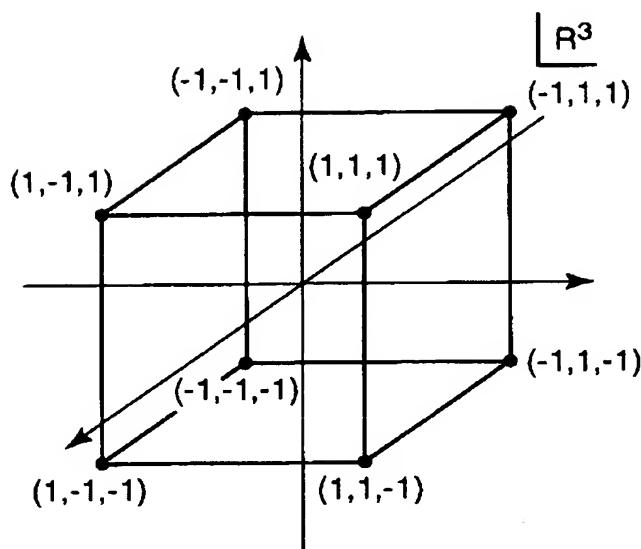
ID	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
S(1)	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
S(2)	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e
S(3)	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a

(b)

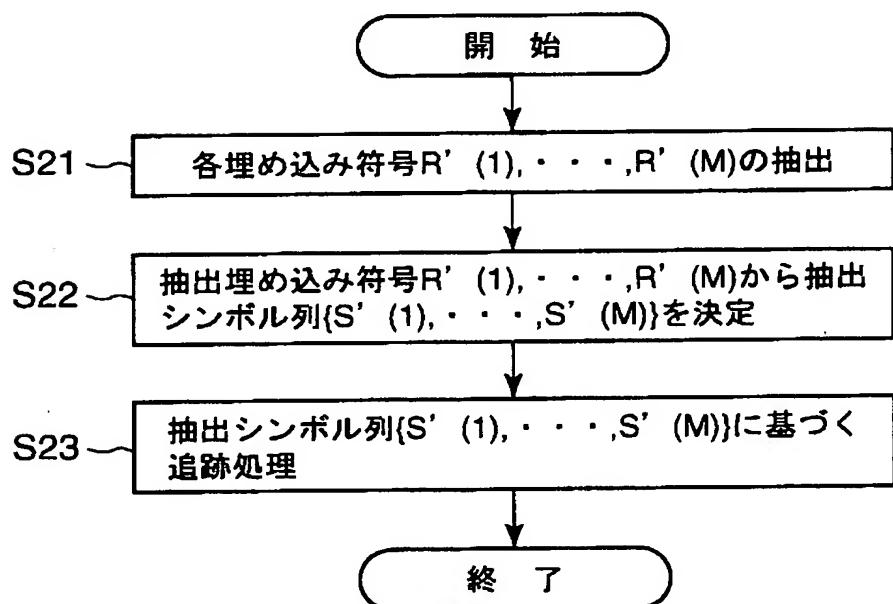
【図6】

ユーザID	ユーザIDに対応するシンボル列	
0	{0,0,0}	{a,a,a}
1	{1,1,1}	{b,b,b}
2	{2,2,2}	{c,c,c}
3	{0,3,3}	{a,d,d}
4	{1,4,4}	{b,e,e}
5	{2,0,5}	{c,a,f}
6	{0,1,6}	{a,b,g}
7	{1,2,0}	{b,c,a}
8	{2,3,1}	{c,d,b}
9	{0,4,2}	{a,e,c}
10	{1,0,3}	{b,a,d}
11	{2,1,4}	{c,b,e}
12	{0,2,5}	{a,c,f}
13	{1,3,6}	{b,d,g}
14	{2,4,0}	{c,e,a}

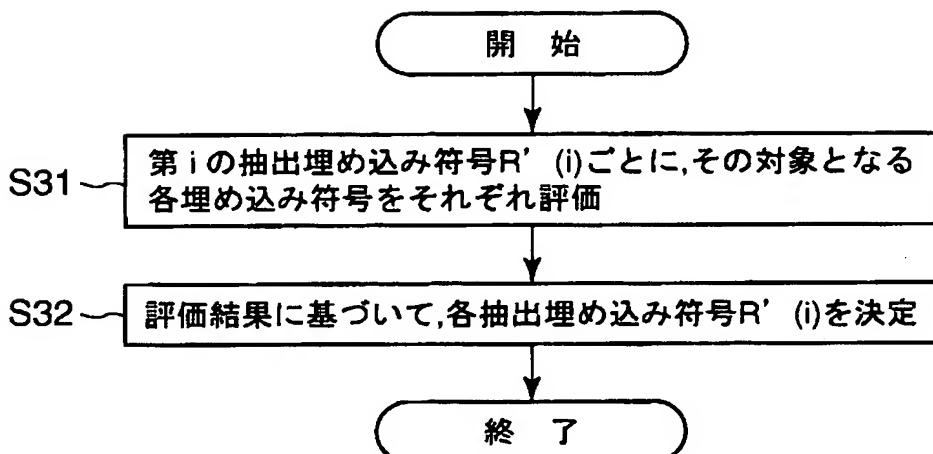
【図7】



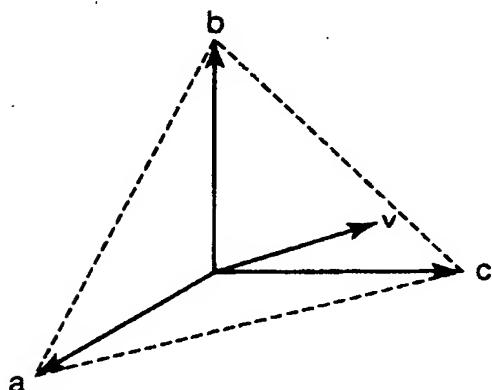
【図8】



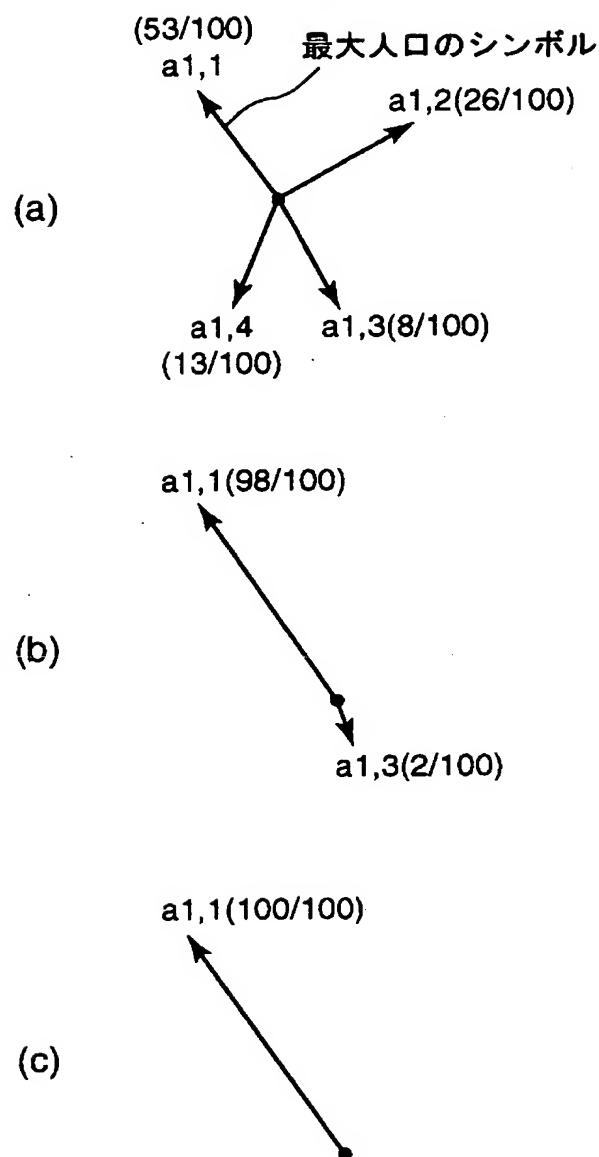
【図9】



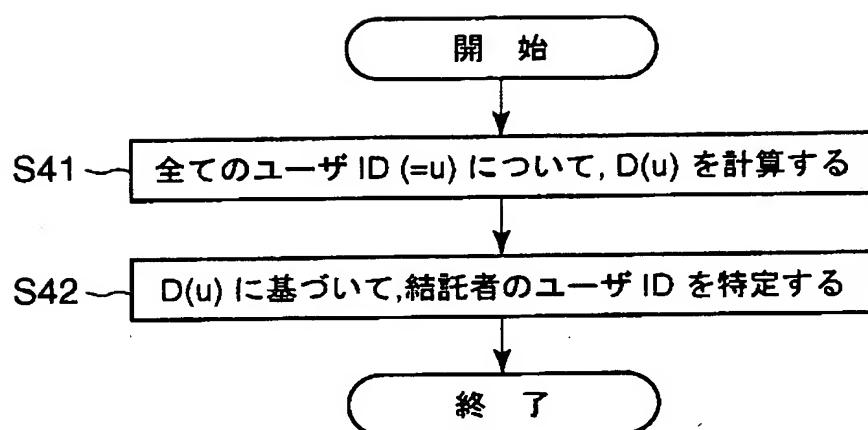
【図10】



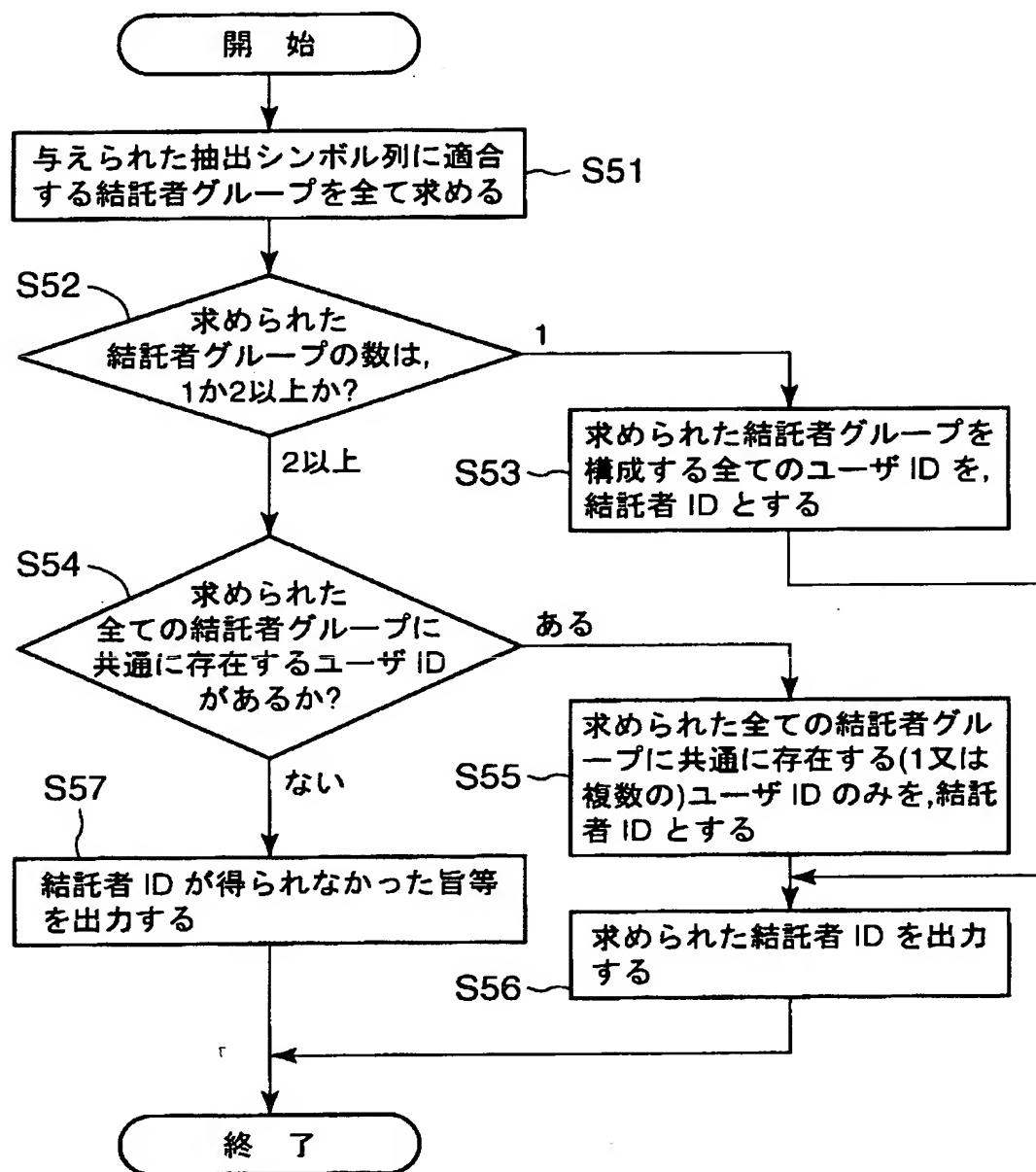
【図11】



【図12】



【図13】



【書類名】 要約書

【要約】

【課題】 結託攻撃に用いられたコンテンツの複製物に埋め込まれていたであろう識別情報を、より感度の高い符号によって推定可能とした電子透かし解析装置を提供すること。

【解決手段】 電子透かし解析装置2は、まず、デジタルコンテンツの複製物から、序列を持った複数の埋め込み符号をそれぞれ抽出する。次いで、抽出された複数の埋め込み符号にそれぞれ対応するシンボルを求め、求められた複数のシンボル列を、対応する埋め込み符号の序列に従って並べてシンボル列を求める。そして、結託攻撃に用いられたデジタルコンテンツの複製物から抽出されたシンボル列及び各々の正当な識別情報に対してそれぞれ一意に割り当てられるシンボル列に基づいて、結託攻撃に用いられたデジタルコンテンツの複製物に埋め込まれていた識別情報を特定する。

【選択図】 図1

出願人履歴情報

識別番号 [000003078]

1. 変更年月日 2001年 7月 2日

[変更理由] 住所変更

住 所 東京都港区芝浦一丁目1番1号

氏 名 株式会社東芝